RL-TR-95-83
Final Technical Report
April 1995

PARAMETRIC SIGNATURE MODELING FOR TARGET IDENTIFICATION

The Ohio State University

Dr. Randolph L. Moses

DTIC
ELECTE
JUL 27 1995
S D
B

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

19950724 007

DTIC QUALITY INSPECTED 5

Rome Laboratory
Air Force Materiel Command
Griffiss Air Force Base, New York

JGK

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RL-TR-95-83 has been reviewed and is approved for publication.

APPROVED: Russell D. Brown

RUSSELL D. BROWN
Project Engineer

FOR THE COMMANDER: Donald W. Hanson

DONALD W. HANSON
Director of Surveillance & Photonics

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL (OCTS) Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1995	Final Aug 91 - Jan 93

4. TITLE AND SUBTITLE

PARAMETRIC SIGNATURE MODELING FOR TARGET IDENTIFICATION

5. FUNDING NUMBERS

C - F30602-91-D-0001,
Task 9
PE - 62702F
PR - 4594
TA - 00
WU - PN

6. AUTHOR(S)

Dr. Randolph L. Moses

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Ohio State University
Columbus OH 43210

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Rome Laboratory (OCTS)
26 Electronic Pky
Griffiss AFB NY 13441-4514

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

RL-TR-95-83

11. SUPPLEMENTARY NOTES

Rome Laboratory Project Engineer: Russell D. Brown/OCTS/(315) 330-4437

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report describes recent advances in the analysis of radar target signatures for the purpose of target identification. These developments include a statistical analysis of TLS-Prony algorithms, the development and statistical analysis of computationally efficient versions of the algorithm based on data decimation, quantitative performance analysis of radar scattering center detection and resolution, and the application of these methods on measured S-band radar data. It is shown that these parametric modeling techniques have two to four times the range resolution of Fourier-based techniques when applied to measured scattering data. The technique directly estimates the location of scattering centers and parameters that characterize each scattering center for target identification.

14. SUBJECT TERMS

Radar signal processing, Algorithms, Target identification

15. NUMBER OF PAGES

134

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

Contents

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

1. Introduction

This report describes recent advances in the analysis of radar target signatures for the purpose of target identification. The goal of the research has been to develop techniques for automatic classification of radar targets by digital processing of measured scattering signals.

Many approaches have been considered for target identification systems. Of these approaches, we have concentrated on processing the scattering data to extract time (range) domain scattering features. In particular, we are interested in the ranges of the dominant scattering centers, and the amplitude (and polarization) properties of each scattering center. These scattering center ranges, along with their associated properties, would form a feature vector for use in a target classification system.

There are two main approaches for extracting the ranges of scattering centers from measured frequency-domain data. Traditionally, target scattering centers are found by using Fourier transform techniques to convert the frequency-domain data into a time-domain, or range-domain, profile [1]. This profile is then searched for sharp peaks, and the ranges of these peaks form the scattering center range estimates. The amplitudes of the peaks give the magnitudes of the scattered field for each scattering center when single polarization data is available. Little work has appeared on Fourier processing of multiple polarization data; one method for extracting polarimetric properties of scattering centers was recently reported in [2].

This report considers a different approach for scattering center extraction: the parametric modeling approach. In parametric modeling, the target scattering characteristics are described by a model, and the parameters of that model are then estimated from measured data. The model is chosen so that the model parameters relate to target scattering centers in a natural way. In addition, the model we have chosen extends naturally to the case in which data for multiple polarizations are available.

An outline of the report is as follows. Section 2 describes the parametric modeling approach in detail, and discusses the advantages of such an approach. Section 3 outlines the technical accomplishments of this research effort. Section 4 concludes the report.

2. Technical Approach

Assume that a given radar acquires stepped frequency measurements of the (coherent) scattering from a target. Denote these measurements as

$$s_{xy}(f_k), \qquad k = 0, 1, \ldots, N-1. \tag{2.1}$$

Here, the subscripts x and y denote the received and transmitted polarizations of the radar, respectively; f_k denotes the frequency at the kth frequency step. For the purpose of clarity, we assume that the radar transmits left-circularly polarized signals and receives both horizontally and vertically polarized fields; thus, both $s_{hl}(f_k)$ and $s_{vl}(f_k)$ are measured. However, other polarizations could be used.

The parametric modeling method involves modeling the target as a set of scattering centers. In particular, we use the Prony, or damped exponential model [3, 4]. In the frequency domain, the model takes the form of sum of damped exponential terms:

$$\begin{bmatrix} s_{hl}(f_k) \\ s_{vl}(f_k) \end{bmatrix} = \sum_{i=1}^{M} \begin{bmatrix} A_{hli} \\ A_{vli} \end{bmatrix} p_i^{f_k}, \qquad 0 \le k \le N-1. \tag{2.2}$$

Here, the A_i and p_i parameters represent the amplitude, range, and damping factor of each of the M scattering centers. Each scattering center is characterized by a pole p_i, whose angle $\angle p_i$ gives the range of the ith scattering center, and whose amplitude $|p_i|$ relates to the range dispersion of the scattering center. For an ideal point scatterer, $|p_i| = 1$, but for realistic targets the scattering will be attenuated slightly as frequency either increases or decreases, thus $|p_i|$ will vary a little bit around one. The complex amplitudes A_{hli} and A_{vli} characterize the amplitude and polarization properties of the ith scattering center; the amplitude, ellipticity, and tilt of the scattering center can be found by a simple transformation of A_{hli} and A_{vli} [3]. This model has the advantage that full polarization information is incorporated into a single model; on

the other hand, if not all of the polarization data is available, the reduced data set can be used to obtain an exponential model with fewer attributes per scattering center.

Parametric modeling of target signatures offers several advantages over traditional Fourier-based target signature analysis. These include:

- **High Resolution:** Parametric methods offer higher resolution than can be obtained by Fourier processing; as a result, lower bandwidth radars may be used to achieve a desired range resolution.

- **Automatic Feature Extraction:** The parametric methods we propose directly estimate the location of scattering centers, and directly estimate a set of parameters which characterize each scattering center. This is in contract to Fourier-based processing techniques, in which some sort of feature extraction procedure must be applied to the estimated range profile of the target.

- **Separate Characterization of Each Scattering Center:** The proposed parametric models characterize each scatterer with a set of attributes which include polarization properties, frequency response, and in some cases a description of the physical nature of that scatterer.

3. Technical Accomplishments

Below we outline the technical accomplishments developed under this research program.

3.1 Statistical Analysis of TLS-Prony Techniques

The main algorithm we used for estimating the parameters of the exponential model is the so-called TLS-Prony estimator. This method involves setting up a set of backward linear equations, and solving for prediction coefficients using a total-least-squares (TLS) techniques. The TLS solution involves performing a singular value decomposition on a data matrix. The zeros of the prediction polynomial associated with the backward linear prediction coefficients form the exponential pole estimates. Once the exponential terms are found, the corresponding amplitude coefficients are then determined by a least-squares technique. This method is attractive because it is computationally simple (involving linear algebraic solutions of equations and a single polynomial rootfinding step), but gives estimates whose statistical properties are close to optimal.

Our first main contribution is a complete statistical analysis of the poles and amplitude coefficients obtained from the TLS-Prony technique when the measurement data are corrupted by Gaussian white noise. The analysis gives the statistics of the model coefficients for high SNR. Because the pole and amplitude parameters relate directly to physical properties of the scattering centers, this analysis gives a theoretical prediction of scattering center extraction performance in the presence of noise (see also [5]). In particular, resolution of scattering centers, and accuracy of scattering amplitudes and polarization properties can be found as a function of SNR and/or polarization diversity. The statistical theory is compared to Monte-Carlo simulation results, and found to agree quite well even for moderate SNRs (-5 dB to 5 dB depending on the signal scenario). These results have been published as a M.S.

Thesis by Mr. Ching-Hui Ying [6], and also as a conference paper presented at the 1991 Asilomar Conference [7]. These results have also been accepted for publication in archival Journal *Automatica* in a special issue on Statistical Signal Processing and Control [8]. The *Automatica* paper is included in Appendix A.

3.2 Computationally Efficient TLS-Prony Estimation Algorithms

A second advance is the development and analysis of TLS-Prony estimation algorithms which have lower computational cost then the standard TLS-Prony method, but which have almost the same accuracy performance. Specifically, the method involves decimation of the frequency-domain scattering measurements. Decimation reduces the length of the data, and gives a large corresponding reduction in computational cost.

Specifically, the data is processed through several different filtering and decimation blocks. Each of these blocks is designed to estimate the scattering centers of the target for a particular segment of the range profile. The individual segments can be processed in parallel, resulting in a significant savings of computation (at an increase of processing hardware). Alternately, the processing can be done sequentially. In either case, the total number of computations is substantially less than the computations involved for a single estimate of all scattering centers.

One concern with such an approach is that the statistical accuracy of the resulting scattering center estimates is lower than the accuracy obtained in a single estimate approach. To address this issue, we derived analytical expressions which predict the accuracy of the scattering center range and amplitude (and, in the case of multiple polarization data, the polarization characteristics of the scattering center as well) for the decimation-based TLS-Prony algorithm. This analysis again assumes high SNR, but is shown through comparisons with Monte-Carlo experiments to hold for moderate and low SNRs as well. The analysis and simulation demonstrate that

the decimation-based algorithm can provide scattering center estimates with almost the same accuracy as in the single estimation algorithm (the loss is about 2-4 dB), but at huge computational savings (reductions by a factor of 20, using sequential computation on a single CPU, were for typical examples; if parallel computation hardware is used, the savings would be greater by an additional factor of about 6–10).

The decimation-based algorithms, and the statistical analysis of these algorithms, have been published in a conference paper which appeared in the 1992 SPIE Aerospace Sensing Symposium [9], and has also been submitted to the IEEE Transactions on Signal Processing [10]. The IEEE Transactions paper is included in Appendix B.

3.3 Detection and Resolution of Scattering Centers

This research is concerned with detection and resolution of scattering centers which are estimated from noisy stepped-frequency measurements of that target. The research is aimed at answering the following questions:

1. As a function of radar bandwidth, SNR, and polarization diversity, how closely spaced can scattering centers be before they can no longer be resolved?

2. What is the probability of detection versus probability of false alarm for a scattering center (as a function of bandwidth, SNR, and polarization diversity)?

The question of resolution is closely related to the statistical analysis of the TLS-Prony methods as described previously. Basically, the resolution question is answered by appropriately applying these analysis results to the radar scattering application.

The detection question arises from the mechanics of the TLS-Prony methods. The TLS-Prony methods achieve high resolution by first finding a backward linear prediction filter of high order. The zeros of this prediction filter are found; some of these zeros correspond to scattering centers on the target, and some are extraneous,

"computation" zeros. One must then decide which zeros correspond to scattering centers; this decision is based on the energy of the "mode" associated with that zero. Because the data are noisy, there will be some times in which a computation zero has energy which exceeds the decision threshold, giving rise to a false alarm scattering center. Similarly, there will be times in which a true scattering center has energy which lies below the threshold, and the scattering center will not be detected. By varying the threshold, the detection versus false alarm probability can be determined for various radar parameter settings (such as SNR, polarization, and bandwidth).

The detection question cannot be answered directly in terms of our previous analyses, because these analyses did not include the statistical properties of the extraneous poles. Thus, we first derived the statistical properties of the extraneous poles; this derivation is presented in [11]. From this analysis, and using properties of the energy of a scattering center, the desired detection performance measures could then be calculated.

A conference paper which describes this work was presented at the International Radar '92 conference, and appears in the conference proceedings [12]. The This paper is included in Appendix C.

3.4 Scattering Center Extraction from Measured Radar Data

Finally, we applied our parametric modeling results to some measured S-band Linear FM scattering data which was supplied to us by Rome Laboratories. Unfortunately, only a limited amount of data was available, and this data was taken at a single polarization, so we were unable to make extensive comparisons between the theoretical statistical analyses and the results obtained from measured radar data. In addition, we had no calibration standard for the data, so it was not possible to compare the estimated scattering centers with a "ground truth". However, we were able to conduct limited studies on scattering center accuracy, and on bandwidth extrapo-

lation capability, of the TLS-Prony algorithms. Based on these studies, we found that scattering center resolutions between 2–4 times the resolution capability of Fourier techniques could be obtained.

The experimental study on measured radar data is presented in detail in the technical report "Prony Modeling of Linear FM Radar Data," [11]; this report is included in Appendix D.

4. Conclusions

This report presents the technical advances in radar scattering center extraction which were developed under this contract. We have made significant advances in the development and analysis of Prony-based scattering center estimation. We have developed theoretical analyses which predict the performance of these algorithms for scattering center resolution, accuracy of scattering center amplitude and polarization characteristics, and detection versus false alarm probabilities of scattering centers. This analysis provides these performance measures as functions of SNR, data length (which is directly related to data bandwidth), and data diversity (which is directly related to polarization diversity of the data). The analysis assumes high SNR, but is shown by Monte-Carlo simulations to be applicable for low to moderate SNRs as well. We have also applied these algorithms to a limited set of measured radar data, and shown scattering center accuracies of two to four times that obtainable with Fourier-based techniques. The results of this research have been published in two technical reports, three conference papers, and two archival journal papers.

Future work should focus on more extensive testing of these methods on measured data, especially fully polarimetric data. In addition, future signal processing research on constrained TLS-Prony or other algorithms (in which, for example, the poles are constrained to lie on or near the unit circle), and on automatic model order selection are of interest.

5. Bibliography

[1] D. R. Wehner, *High Resolution Radar.* Dedham, MA: Artech House, 1987.

[2] N. Chamberlain, E. K. Walton, and F. D. Garber, "Radar target identification of aircraft using polarization-diverse features," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-27, pp. 58–67, January 1991.

[3] W. M. Steedly and R. L. Moses, "High resolution exponential modeling of fully polarized radar returns," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-27, pp. 459–469, May 1991.

[4] R. Carrière and R. L. Moses, "High resolution radar target modeling using a modified Prony estimator," *IEEE Trans. Antennas and Propagation*, vol. AP-40, pp. 13–18, January 1992.

[5] C. J. Ying, R. L. Moses, and R. L. Dilsavor, "Perturbation Analysis for Pole Estimates of Damped Exponential Signals," technical report, The Ohio State University, Department of Electrical Engineering, ElectroScience Laboratory, August 1991.

[6] C.-H. J. Ying, "Statistical analysis of TLS-based Prony techniques," Master's thesis, The Ohio State University, March 1992.

[7] W. M. Steedly, C. J. Ying, and R. L. Moses, "Statistical Analysis of SVD-Based Prony Techniques," in *Proceedings of the Twenty-Fifth Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA), pp. 232–236, November 4–6 1991.

[8] W. M. Steedly, C. J. Ying, and R. L. Moses, "Statistical analysis of TLS-based Prony techniques," *Automatica, Special Issue on Statistical Signal Processing and Control.* Accepted.

[9] W. M. Steedly, C. J. Ying, and R. L. Moses, "A modified TLS-Prony method using data decimation," in *Proceedings of the SPIE OE/AEROSPACE Science and Sensing Conference*, (Orlando, FL), April 22–24, 1992.

[10] W. M. Steedly, C. J. Ying, and R. L. Moses, "A Modified TLS-Prony Method Using Data Decimation for Computationally Efficient Estimation of Damped Exponentials in Noise," *IEEE Transactions on Signal Processing.* submitted.

[11] C. J. Ying, W. M. Steedly, and R. L. Moses, "Statistical Analysis of True and Extraneous TLS-Prony Mode Estimates," SPANN Laboratory Technical Report TR-92-02, The Ohio State University, July 1992.

[12] W. M. Steedly, C. J. Ying, and R. L. Moses, "Resolution bound and detection results for scattering centers," in *Proceedings of the Int'l Radar 92 Conference*, (Brighton, UK), Oct 12–13 1992. (to appear).

A. Preprint of "Statistical Analysis of TLS-Based Prony Techniques"

The following pages contain a preprint of the paper, "Statistical Analysis of TLS-Based Prony Techniques," which is accepted for publication in *Automatica, Special Issue on Statistical Signal Processing and Control.*

Statistical Analysis of TLS-Based Prony Techniques [†]

William M. Steedly
The Analytic Sciences Corporation
Reston, VA 22090

Ching-Hui J. Ying and Randolph L. Moses
Department of Electrical Engineering
The Ohio State University
Columbus, OH 43210

Abstract

We present an analysis of parameter variance statistics for the TLS-Prony method applied to damped exponential signals. We derive the covariance matrix of the estimated parameters for this method. The parameters include the magnitudes and angles of the poles, and the magnitudes and angles of the amplitude coefficients. We verify the theoretical results using Monte-Carlo simulations studies. We also compare the variance results to the corresponding Cramér-Rao bounds for several cases.

Submitted to Automatica
Special Issue on Statistical Signal Processing and Control
November 1991
Revised August 1992

[†]This research was supported by the Air Force Office of Scientific Research, the Avionics Division, Wright Laboratories, and the Surveillance Division, Rome Laboratories.

I. Introduction

The problem of estimating model parameters of noisy exponential signals is an active area of research. This problem has applications in a number of areas, including speech processing, deconvolution, radar and sonar signal processing, array processing, and spectrum estimation. A number of authors have considered various aspects of this problem [1]–[22], and a large number of algorithms have been developed and analyzed [8, 11, 12], [16]–[18], [20, 21].

One popular class of algorithms for estimating parameters from noisy exponential sequences are the subspace-based approaches [1]–[10], [13]–[16], [19]–[22]. These include the MUSIC algorithm and its enhancements [1, 16, 21], subspace rotation methods such as ESPRIT [5, 14, 15, 20, 21], iterative maximum likelihood methods [6, 7, 13], minimum norm methods [8, 19], and total least squares (TLS) methods [2, 9, 10, 22]. These methods have proven attractive because they exhibit good statistical performance at a modest computational cost. This has been well-established by a large number of numerical studies.

More recently, there has been interest in quantitatively evaluating these methods. To this end, a number of researchers have analyzed the statistical properties of such algorithms [3, 8, 11, 12, 16, 17, 18, 20, 21]. In [3], Henderson presents a geometric study of the pole estimation problem, and analyzes the statistical properties of the prediction coefficients when the Hankel data matrix is corrupted by an i.i.d. noise matrix. Several authors have presented results relating to pole angle (frequency) estimates from arrays when the exponential signals are undamped, *e.g.*, [8, 11, 16, 17, 20, 21, 23]. A related perturbation analysis of SVD-based methods is presented

in [24, 25, 26, 27] and applied to both frequency estimation and threshold analysis for exponential modes. In [12] Hua and Sarkar present an angle-only analysis for the least squares Prony method for the poles of undamped exponentials. Less has appeared which considers the statistical properties of the parameters for damped exponentials. Porat and Friedlander [28] consider the related problem of ARMA system identification using SVD-based approaches. Hua and Sarkar [18] present an analysis for the pole estimates of damped exponential signals using their matrix pencil method, but have not presented the statistical properties of the amplitude coefficients.

This paper presents an extension of the above works to treat a general exponential case. We introduce a complete statistical derivation for the poles and amplitude coefficients estimated using a TLS-Prony scheme where signals consist of arbitrary damped exponential terms in noise. We provide complete statistics for the individual pole parameters for an exponential model in which the poles may lie on, inside, or outside the unit circle. In addition, we derive the statistical properties of the amplitude coefficients associated with these exponential modes.

The results of this paper provide a sound basis for performance analyses of the TLS-Prony estimation method. We extend previous works by considering the general damped case, as well as by including amplitude coefficient parameters in addition to pole parameters. These results provide the tools to analyze various situations and evaluate the potential success of applying the TLS-Prony estimation algorithm.

The TLS-Prony estimation procedure is a multi-snapshot extension of the algorithm presented in [9, 10]. The advantage of singular value decomposition (SVD) in noise cleaning of the Toeplitz data matrix is well-known. The multiple snapshot incor-

poration is a straightforward one in which more than one set of amplitude coefficients corresponds to the set of poles. The procedure is discussed in Section II.

The statistical derivation for this procedure is based on a first order perturbation analysis; thus the analysis assumes high SNR. We derive the complete covariance matrix of the estimated parameters for this case. The parameters include the magnitudes and angles of the poles, and the magnitudes and angles of the amplitude coefficients.

Using these expressions, several general properties of the parameter covariance matrix are derived for the high SNR case. We show that the angle and magnitude parameters are uncorrelated for both the poles and the amplitude coefficients. We also show that if the relative magnitude of the pole or amplitude coefficient estimate is considered (*i.e.* $\frac{\widehat{\alpha}}{\alpha}$, where α is the true magnitude), then the corresponding angle and relative magnitude variances are equal.

This paper also examines pole estimation accuracy as functions of pole magnitude, data length, and pole separation using the variance expressions. We compare these variance results to the corresponding Cramér-Rao bounds and verify the theoretical results using Monte-Carlo simulations. The effects on poles inside and outside the unit circle using backward or forward linear prediction in the TLS-Prony estimation scheme are also detailed.

An outline of this paper is as follows. Section II presents the data model. Section III presents the statistics of the model parameters and their properties. Section IV presents some examples using the statistical expressions. Finally, Section V concludes the paper.

II. Estimation Procedure Review

A. *Data Model*

Assume we have N "snapshots" of data vectors $y(t)$, each of length m:

$$y(t) = \begin{bmatrix} y_0(t) & y_1(t) & \cdots & y_{m-1}(t) \end{bmatrix}^T \qquad t = 1, 2, \ldots, N. \tag{1}$$

Each data vector is modeled as a noisy exponential sequence

$$y_q(t) = \sum_{i=1}^{n} x_i(t) p_i^q + e_q(t) \qquad q = 0, 1, \ldots, m-1. \tag{2}$$

There are n distinct exponential modes in the data; the n poles $\{p_i\}_{i=1}^n$ do not vary from snapshot to snapshot, but the amplitudes $x_i(t)$ may vary. Here, it is assumed that $\{e_q(t)\}$ are uncorrelated zero mean complex white Gaussian noise sequences with variance σ. Equation (2) may be compactly written as

$$y(t) = Ax(t) + e(t), \tag{3}$$

where $e(t) = \begin{bmatrix} e_0(t) & e_1(t) & \cdots & e_{m-1}(t) \end{bmatrix}^T$, $x(t) = \begin{bmatrix} x_1(t) & x_2(t) & \cdots & x_n(t) \end{bmatrix}^T$, and A is the $m \times n$ Vandermonde matrix derived from n signal poles

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ p_1 & p_2 & \cdots & p_n \\ p_1^2 & p_2^2 & \cdots & p_n^2 \\ \vdots & \vdots & & \vdots \\ p_1^{m-1} & p_2^{m-1} & \cdots & p_n^{m-1} \end{bmatrix}. \tag{4}$$

B. Parameter Estimation

The multi-snapshot backward linear prediction equations are given by:

$$\left[\begin{array}{c:c} y & Y \end{array}\right]\left[\begin{array}{c} 1 \\ \hline b \end{array}\right] \approx 0, \tag{5}$$

where

$$b = \left[\begin{array}{cccc} b_1 & b_2 & \cdots & b_L \end{array}\right]^T \tag{6}$$

and where

$$\left[\begin{array}{c:c} y & Y \end{array}\right] = \left[\begin{array}{ccccc} y_0(1) & y_1(1) & y_2(1) & \cdots & y_L(1) \\ y_1(1) & y_2(1) & y_3(1) & \cdots & y_{L+1}(1) \\ \vdots & \vdots & \vdots & & \vdots \\ y_{m-(L+1)}(1) & y_{m-L}(1) & y_{m-(L-1)}(1) & \cdots & y_{m-1}(1) \\ \hline y_0(2) & y_1(2) & y_2(2) & \cdots & y_L(2) \\ y_1(2) & y_2(2) & y_3(2) & \cdots & y_{L+1}(2) \\ \vdots & \vdots & \vdots & & \vdots \\ y_{m-(L+1)}(2) & y_{m-L}(2) & y_{m-(L-1)}(2) & \cdots & y_{m-1}(2) \\ \hline & & \vdots & & \\ \hline y_0(N) & y_1(N) & y_2(N) & \cdots & y_L(N) \\ y_1(N) & y_2(N) & y_3(N) & \cdots & y_{L+1}(N) \\ \vdots & \vdots & \vdots & & \vdots \\ y_{m-(L+1)}(N) & y_{m-L}(N) & y_{m-(L-1)}(N) & \cdots & y_{m-1}(N) \end{array}\right]. \tag{7}$$

Here L is the order of prediction and b is the coefficient vector of the polynomial $B(z)$ given by

$$B(z) = 1 + b_1 z^1 + b_2 z^2 + \cdots + b_L z^L. \tag{8}$$

For the noiseless case, L can be any integer greater than or equal to the model order n; however, in the presence of noise choosing $L > n$ results in more accurate parameter

estimates (see Section II). Note that all of the N snapshots are used simultaneously to estimate a single set of prediction coefficients (and therefore, a single set of poles).

The TLS-Prony method considers the effect of noise perturbation of both Y and y, and the TLS solution attempts to minimize the effect of these perturbations on the prediction coefficient vector b (see [9, 10] for details). This is accomplished by obtaining an SVD of the matrix $\left[\begin{array}{c:c} y & Y \end{array}\right]$ and truncating all but the first n singular values to arrive at an estimate $\left[\begin{array}{c:c} \hat{y} & \hat{Y} \end{array}\right]$ [9, 10]. Inserting $\left[\begin{array}{c:c} \hat{y} & \hat{Y} \end{array}\right]$ in Equation (5) gives the modified linear prediction equation

$$\hat{Y}\hat{b} = -\hat{y} \tag{9}$$

from which the linear prediction coefficient vector estimate $\hat{b}$ is found as

$$\hat{b} = -\hat{Y}^{+}\hat{y}, \tag{10}$$

where $^{+}$ denotes the Moore-Penrose pseudoinverse. A numerically robust solution for $\hat{b}$ can be found directly from the SVD of $\left[\begin{array}{c:c} \hat{y} & widehatY \end{array}\right]$, as is shown in [10]. Finally, the estimates for the poles are found by

$$\hat{p}_j = \text{zero}_j\left(\hat{B}(z)\right), \qquad j = 1, 2, \ldots, L. \tag{11}$$

Once the L poles are determined from Equation (11), one must separate the n "true" poles from the remaining $L - n$ "extraneous" or "noise" poles. A popular approach is to choose n poles based on their location with respect to the unit circle. For example, one can choose the n poles closest to the unit circle if it is known that the poles are undamped [29] or the n poles with smallest moduli if it is known that the poles are damped [4]. However, these methods do not apply when the true

poles may lie both inside and outside the unit circle. In this case we can classify poles as true or extraneous based on the energy of the corresponding mode. We have found this method to be more reliable than other procedures for the case when the true modes lie both inside and outside the unit circle. This arises, for example, in the radar scattering problem where measurements are made over a small relative bandwidth, and the exponential modes in the data can be decaying or growing over that band [30, 31].

In this energy criterion method, the L sets of amplitude coefficients can be found using the pole estimates, and Equation (3) leads to the following least squares equation for the amplitude coefficients,

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \hat{p}_1 & \hat{p}_2 & \cdots & \hat{p}_L \\ \hat{p}_1^2 & \hat{p}_2^2 & \cdots & \hat{p}_L^2 \\ \vdots & \vdots & & \vdots \\ \hat{p}_1^{n-1} & \hat{p}_2^{n-1} & \cdots & \hat{p}_L^{n-1} \end{bmatrix} \begin{bmatrix} \hat{x}_1(1) & \hat{x}_1(2) & \cdots & \hat{x}_1(N) \\ \hat{x}_2(1) & \hat{x}_2(2) & \cdots & \hat{x}_2(N) \\ \vdots & \vdots & & \vdots \\ \hat{x}_L(1) & \hat{x}_L(2) & \cdots & \hat{x}_L(N) \end{bmatrix} \approx \begin{bmatrix} y(1) & y(2) & \cdots & y(N) \end{bmatrix} \tag{12}$$

or

$$\hat{A}_L \hat{X}_L \approx Y_a. \tag{13}$$

A least squares solution to Equation (13) can be computed as

$$\hat{X}_L = \left(\hat{A}_L^* \hat{A}_L\right)^{-1} \hat{A}_L^* Y_a = \hat{A}_L^\dagger Y_a, \tag{14}$$

where * denotes complex conjugate transpose. (However, in practice, more numerically robust procedures, such as a QR decomposition, should be used to solve Equation (13).) Because only n singular values of $\begin{bmatrix} \hat{y} & : & \hat{Y} \end{bmatrix}$ are nonzero, there are at

most n pole estimates which can correspond to data modes. Therefore, only the n poles which have the largest energy are retained. This is done by computing the L mode energies as

$$E_i = \sum_{t=1}^{N} |\hat{x}_i(t)|^2 \sum_{q=0}^{m-1} |\hat{p}_i|^{2q} \qquad i = 1, 2, \ldots, L \tag{15}$$

and retaining those n poles whose corresponding energies are highest. We then reestimate the amplitude coefficients of these n poles. This is done using Equation (14), except that $\hat{A}_L$ is replaced by $\hat{A}$, where $\hat{A}$ is the Vandermonde matrix composed only of the n columns of $\hat{A}_L$ corresponding to the n selected poles. We note that the above procedure produces consistent estimates as $\sigma \to 0$, as is shown in the Appendix.

III. Statistical Analysis

In this section we present a first order statistics of the complete set of parameter estimates obtained in the TLS-Prony method. Assume the data is given as in Equation (2). Let ω_i and α_i be the angle and magnitude, respectively, of each pole p_i, thus $p_i = \alpha_i e^{j\omega_i}$. Similarly let $\gamma(t)$ and $\beta(t)$ be the angle and magnitude vectors, respectively, of each vector of amplitude coefficients $x(t)$,

$$\begin{aligned} \gamma(t) &= \begin{bmatrix} \gamma_1(t) & \gamma_2(t) & \cdots & \gamma_n(t) \end{bmatrix}^T \\ \beta(t) &= \begin{bmatrix} \beta_1(t) & \beta_2(t) & \cdots & \beta_n(t) \end{bmatrix}^T, \end{aligned} \tag{16}$$

where

$$x(t) = \begin{bmatrix} \beta_1(t)e^{j\gamma_1(t)} & \beta_2(t)e^{j\gamma_2(t)} & \cdots & \beta_n(t)e^{j\gamma_n(t)} \end{bmatrix}^T. \tag{17}$$

Define following parameter vectors:

$$\begin{aligned}
\theta_x &= \left[\begin{array}{cccccccc} \gamma^T(1) & \beta^T(1) & \gamma^T(2) & \beta^T(2) & \cdots & \gamma^T(N) & \beta^T(N) \end{array} \right]^T \\
\theta_p &= \left[\begin{array}{cccccccc} \omega_1 & \omega_2 & \cdots & \omega_n & \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{array} \right]^T \\
\theta &= \left[\begin{array}{cc} \theta_x^T & \theta_p^T \end{array} \right]^T .
\end{aligned} \tag{18}$$

The following theorem gives the first order approximation of the probability density function (pdf) of $\hat{\theta}$.

Theorem 1: Let $\hat{\theta}$ denote the TLS-Prony estimate of θ which is given by the n highest energy mode estimates found in Equations (11) and (14). Then to a first order approximation (as $\sigma \to 0$), the pdf of $\hat{\theta}$ is given by the unbiased Normal distribution

$$\hat{\theta} \sim \mathcal{N}(\theta, \Sigma_\theta), \tag{19}$$

where

$$\Sigma_\theta = \left[\begin{array}{cccccc}
\overline{U}(1,1) & \widetilde{U}(1,1)T_\beta^{-1}(1) & \cdots & \overline{U}(1,N) & \widetilde{U}(1,N)T_\beta^{-1}(N) & \overline{V}(1) & \widetilde{V}(1)T_\alpha^{-1} \\
-T_\beta^{-1}(1)\widetilde{U}(1,1) & T_\beta^{-1}(1)\overline{U}(1,1)T_\beta^{-1}(1) & \cdots & -T_\beta^{-1}(1)\widetilde{U}(1,N) & T_\beta^{-1}(1)\overline{U}(1,N)T_\beta^{-1}(N) & -T_\beta^{-1}(1)\widetilde{V}(1) & T_\beta^{-1}(1)\overline{V}(1)T_\alpha^{-1} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\overline{U}(N,1) & \widetilde{U}(N,1)T_\beta^{-1}(1) & \cdots & \overline{U}(N,N) & \widetilde{U}(N,N)T_\beta^{-1}(N) & \overline{V}(N) & \widetilde{V}(N)T_\alpha^{-1} \\
-T_\beta^{-1}(N)\widetilde{U}(N,1) & T_\beta^{-1}(N)\overline{U}(N,1)T_\beta^{-1}(1) & \cdots & -T_\beta^{-1}(N)\widetilde{U}(N,N) & T_\beta^{-1}(N)\overline{U}(N,N)T_\beta^{-1}(N) & -T_\beta^{-1}(N)\widetilde{V}(N) & T_\beta^{-1}(N)\overline{V}(N)T_\alpha^{-1} \\
\overline{V^*}(1) & \widetilde{V^*}(1)T_\beta^{-1}(1) & \cdots & \overline{V^*}(N) & \widetilde{V^*}(N)T_\beta^{-1}(N) & \overline{Z} & \widetilde{Z}T_\alpha^{-1} \\
-T_\alpha^{-1}\widetilde{V^*}(1) & T_\alpha^{-1}\overline{V^*}(1)T_\beta^{-1}(1) & \cdots & -T_\alpha^{-1}\widetilde{V^*}(N) & T_\alpha^{-1}\overline{V^*}(N)T_\beta^{-1}(N) & -T_\alpha^{-1}\widetilde{Z^*} & T_\alpha^{-1}\overline{Z}T_\alpha^{-1}
\end{array} \right], \tag{20}$$

where $\bar{\cdot}$ and $\tilde{\cdot}$ in Equation (20) are real and imaginary part operators, respectively, and where $U(t,r)$, $V(t)$, and Z are $n \times n$ complex matrices which depend on θ, L, and m. The specific formulas for these entries can be found in the Appendix. $T_\beta(t)$

and T_α are diagonal matrices given by

$$\begin{aligned} T_\beta(t) &= \text{diag}\left(\frac{1}{\beta_1(t)}, \frac{1}{\beta_2(t)}, \ldots, \frac{1}{\beta_n(t)}\right) \\ T_\alpha &= \text{diag}\left(\frac{1}{\alpha_1}, \frac{1}{\alpha_2}, \ldots, \frac{1}{\alpha_n}\right). \end{aligned} \tag{21}$$

Proof: See the Appendix. □

Several properties of the covariance can be derived from the structure of Σ_θ. Some of these properties are presented in the following corollaries.

Corollary 1: From Σ_θ in Equation (20), $\text{Cov}\left(\widehat{\gamma_i}(t), \widehat{\beta_i}(t)\right) = 0$, and $\text{Cov}\left(\widehat{\omega_i}, \widehat{\alpha_i}\right) = 0$.

Proof: Consider the blocks of Σ_θ containing the covariances of interest, which are given by $\tilde{U}(t,t){T_\beta}^{-1}(t)$ and $\tilde{Z}{T_\alpha}^{-1}$. From Equation (66) in the Appendix, it can be seen that $U(t,t)$ and Z are Hermitian. If follows that the diagonal elements of $\tilde{U}(t,t)$ and $\tilde{Z}$ are zero. Since $T_\beta(t)$ and T_α are real, diagonal matrices, the diagonal elements of $\tilde{U}(t,t){T_\beta}^{-1}(t)$ and $\tilde{Z}{T_\alpha}^{-1}$ are also zero, which gives the desired result. □

Note that when $t \neq r$, $U(t,r)$ is not Hermitian, so the diagonal elements of $\tilde{U}(t,r)$ are not zero. Thus it is not in general the case that the magnitude of $x_i(t)$ and the angle of $x_j(t)$ are uncorrelated for $i \neq j$.

Note that from Equation (20) the angle variances are equal to the magnitude variances except for the transformation matrices $T_\beta(t)$ and T_α. These transformation matrices can be eliminated by rescaling some of the parameters in θ. The required rescaling is obtained by using the *relative* magnitudes of the poles and amplitude coefficients as the estimated parameters instead of their absolute magnitudes. That is, define the estimate $\hat{\theta}_1$ to be as in Equation (18), but with $\hat{\alpha}_i$ and $\hat{\beta}_i(t)$ replaced by

the relative magnitudes $\frac{\widehat{\alpha}_i}{\alpha_i}$ and $\frac{\widehat{\beta}_i(t)}{\beta_i(t)}$. We then note that the Jacobian transformation from θ to θ_1 is given by

$$J = \text{diag}\left(I_n, T_\beta(1), I_n, T_\beta(2), \ldots, I_n, T_\beta(N), I_n, T_\alpha\right). \tag{22}$$

Corollary 2: $\Sigma_{\theta_1} = \text{cov}\left(\widehat{\theta}_1\right)$ is given by Equation (20) with all $T_\beta(t)$ and T_α matrices replaced by identity matrices. It follows that the covariances of parameter angles are equal to the covariances of the corresponding relative magnitudes.

Proof: Immediate from the fact that $\Sigma_{\theta_1} = J\Sigma_\theta J$ with J defined in Equation (22). □

We can also consider a reparameterization of θ in which real and imaginary parts of the amplitude coefficients and poles are considered as parameters. Let us denote such a reparameterization as θ_2, with corresponding covariance matrix which would give Σ_{θ_2}.

Corollary 3: Let ν denote a complex parameter, which is either a pole p_i or an amplitude coefficient $x_i(t)$. Then $\text{var}(\text{Re}\{\nu\}) = \text{var}(\text{Im}\{\nu\})$ and $\text{Re}\{\nu\}$ is uncorrelated with $\text{Im}\{\nu\}$.

Proof: The result can be obtained by applying the Jacobian variable transformation from polar to rectangular coordinates to Σ_θ. This transformation is straightforward, but tedious, and not presented here. □

Corollary 4: Σ_θ is independent of the absolute phase references of the amplitude coefficients within each snapshot, $\phi(t)$, and independent of the absolute phase reference of the poles, ϕ.

Proof: The result follows by examining the expressions for $T_\beta(t)$ and T_α in Equation (21), $U(t,r)$, $V(t)$, and Z in the Appendix, and noting that they remain un-

changed if $T_x(t)$ is replaced by $e^{j\phi(t)}T_x(t)$ and p_i is replaced by $e^{j\phi}p_i$. □

IV. Examples

In this section we present a set of examples which illustrate the performance of the TLS-Prony method. We first compare the first order statistics presented above to the CRB for a number of cases. The CRB for this data model is presented in [32]. We then compare the first order statistics to those obtained using Monte-Carlo simulation.

A. Example 1: Single Exponential Mode

In this example we consider a single pole model with one snapshot of data (and thus one amplitude coefficient). The experiment entails moving the pole along the positive real axis from 0.1 to 10, *i.e.* $0.1 \leq p \leq 10$ (the results are independent of the pole angle by Corollary 4, so an angle of zero is chosen). For each pole location , we calculate the parameter variances using Equation (20) for data sets of lengths 2, 5, 10, 20, 50, and 100. For comparative purposes, the amplitude coefficient associated with the pole is chosen to be a positive real number such that the mode energy $(x^2 \sum_{l=0}^{m-1} p^{2l})$ is unity for each pole location and data length. The noise power is also kept constant at $\sigma = 1$. The model order L is chosen to be one third of the data length m, which has been shown to be near optimal for a number of cases (see [18, 33], and *A*.3 below).

1. Pole Variances

The first order theoretical variances for the estimated pole angle and pole magnitude appear as the dashed lines in Figures 1 and 2, respectively; the corresponding CRBs appear as the solid lines in these figures. From Figure 1 we see that the pole angle variances are close to the CRBs when the pole is inside the unit circle. When the pole is outside the unit circle, the variances become much higher than the CRBs (except for the $m = 2$ case). For larger data lengths the disparity with the CRB is much more pronounced. This is because backward linear prediction is used in our TLS-Prony estimation method. With backward linear prediction, extraneous poles lie outside the unit circle, thus making estimation of poles outside the unit circle more difficult [34]. The use of forward linear prediction would give corresponding results for poles inside the unit circle. Similar observations apply to the pole magnitude variances (see Figure 2). The pole magnitude variances can be normalized to give relative error of the pole magnitude, *i.e.* $\mathrm{var}\left(\frac{\hat{\alpha}_1}{\alpha_1}\right)$. If this is done, one obtains exactly the same curves as in Figure 1 (*cf.* Corollary 2).

From these two figures we also see that inside the unit circle the variances for pole angle are higher than the variances for pole magnitude and vice-versa outside the unit circle. This is because the angular uncertainty becomes greater as a pole moves closer to the origin. As expected, the pole angle variance approaches infinity as the pole approaches the origin.

From Figures 1 and 2 we see that the pole angle and magnitude variances are asymptotically (as $m \to \infty$) lowest when the pole is on the unit circle, and that on the unit circle the variances are decreasing by $1/m^2$ (m is the data length). This is

consistent with the well-known $1/m^3$ variance decrease, since the amplitude coefficient is adjusted in this experiment to keep the mode energy constant (if the amplitude coefficient is left unchanged, the variance decrease is $1/m^3$).

When the pole is not on the unit circle, the variances do not decrease to zero as $m \to \infty$. Recall that we keep the total mode energy constant. For a decaying or growing exponential mode, adding data points while keeping the energy constant results in adding data points with smaller and smaller amplitude. As a result, the parameter estimate variances do not continue to decrease.

2. Amplitude Coefficient Variances

The variances for the amplitude coefficient angle and magnitude appear in Figures 3 and 4. As before, each curve is a plot of variance versus pole radius for various data lengths m. There are several points to note in these figures. First, when the pole is inside the unit circle, increasing the number of data points provides no significant decrease in the variances. The first data point with no noise added is precisely the amplitude coefficient. When the pole is inside the unit circle, this amplitude does not change much as a function of data length, and consequently its variance does not change much. However, when the pole is outside the unit circle, the amplitude coefficient decreases rapidly toward zero as data length increases. Thus, outside the unit circle the estimate of the amplitude cannot be expected to vary much around zero and the magnitude variances become low. Also, variance of the estimated angle becomes quite large because of high angular uncertainty for points near zero.

3. Prediction Order Considerations

In the above figures we used a prediction order L equal to one-third the data length. We next consider the effect of prediction order on the variances of the TLS-Prony parameter estimates. We consider a single exponential mode whose pole is on the unit circle. We choose $\sigma = 1$ and choose the amplitude coefficient so that the mode energy is unity, as before.

Figures 5, 6, and 7 show the variances of the pole angle, amplitude coefficient magnitude, and amplitude coefficient angle. The solid lines represent the CRBs (which are, of course, independent of TLS prediction order), and the dotted lines represent the TLS-Prony variances. Figure 5 has been presented in earlier works [18, 23, 25], but the amplitude coefficient was not considered there. Since the pole is on the unit circle, the pole magnitude results are identical to the pole angle results (*cf.* Corollary 2, with $\alpha_1 = 1$). From these figures we can see that the best prediction order choice is approximately $m/3$; this agrees with the results in [18, 23, 25].

B. Example 2: Two Undamped Exponential Modes

In this example we consider two poles at $\alpha_1 e^{j(\omega_0 + \Delta\omega/2)}$ and $\alpha_2 e^{j(\omega_0 - \Delta\omega/2)}$ with $\alpha_1 = \alpha_2 = 1$. Variances are computed for various data lengths ($m = 5$, 10, 20, 50, and 100) and angle separations $\Delta\omega$. The variance results are independent of ω_0 so $\omega_0 = 0$ is used. Again, $L = m/3$, $\sigma = 1$, one snapshot of data is used ($N = 1$), and each amplitude is chosen such that the corresponding mode energy is unity.

Figure 8 shows the variances for the pole angle estimate (of either pole) as a function of pole separation and data length. The solid lines again show corresponding

CRB results. The variances for the pole magnitudes are equal to the pole angle variances because these poles are located on the unit circle. We see that the TLS-Prony algorithm performs well with respect to the CRB. In fact, for the curves shown, the TLS-Prony variance curves are always within 2dB of the corresponding CRB curves.

C. Example 3: Monte-Carlo Simulation of Ten Mode Case

In this example, we have chosen ten exponential modes to represent a general case. The true pole location of each mode is denoted with an "x" in Figure 9(a). For this case we have $m = 100$ data points, $L = 33$, and $\sigma = 0.001$. The amplitude coefficients are chosen so that each mode energy is unity; this corresponds to an SNR of 10dB per mode; the total SNR (signal power/noise power) is 20dB. The modes were chosen as a representative sampling of various situations.

Figure 9 presents a comparison between the TLS-Prony estimate theoretical variances and variance estimates obtained using Monte-Carlo simulations. The theoretical variances are shown as two-standard deviation concentration ellipses around each pole. These ellipses (they are actually circles, by Corollary 3) give 87% confidence intervals in the complex plane for pole pole estimates. The dots in Figure 9 are pole estimates corresponding to the ten highest energy poles from each of 100 Monte-Carlo simulations. The details of the pole estimation accuracy are summarized in Table 1. Note that the Monte-Carlo variances are close to those predicted by theory for most of the poles, in particular for those closer to the unit circle. For poles well inside the unit circle, there is some bias present which is not predicted by a first order approximation; in addition, the predicted variance is smaller than the actual variance. As poles

Table 1: Theoretical and simulation variances for the poles of a general ten mode case.

Pole number	Theoretical variance (dB)	Simulation variance (dB)	Pole radius
1	-24.2	-22.4	1.115
2	-42.7	-43.4	1.05
3	-35.7	-35.2	0.8
4	-25.8	-22.4	1.12
5	-57.1	-56.9	1.0
6	-57.0	-56.1	1.0
7	-25.8	-21.5	0.6
8	-29.4	-29.1	1.115
9	-39.1	-32.2	0.9
10	-30.1	-26.7	0.7

move outside the unit circle to the radius of the extraneous poles, some deterioration occurs in terms of misidentifying pole estimates as "true" or "extraneous". Note the rapid increase in variance of a pole estimate as its radius increases, by comparing the variance for pole 2, 8, 4, and 1.

V. Conclusions

We have presented a statistical analysis for estimated poles of the TLS-Prony algorithm. This analysis includes complete expressions for the covariance matrix of the parameters of an exponential model which contains one set of poles and multiple sets of amplitude coefficients. The poles of this model may lie anywhere in the complex plane. Using these expressions, several useful properties of the covariance matrix were established. These include independence of the two parameters for each amplitude coefficient and pole, whether one considers a polar, a relative magnitude polar, or a rectangular real and imaginary part parameterization. It was also established that the variances of these pairs are equal for the relative magnitude polar and rectangular real and imaginary part parameterizations.

The results of this paper provide a sound basis for performance analyses of the TLS-Prony estimation method. We have extended previous works to include the general damped undamped case, as well as to include amplitude coefficient parameters in addition to pole parameters. The results can be used to analyze various situations and evaluate the potential success of applying the TLS-Prony estimation algorithm, as the corollaries and examples in the paper demonstrate.

Appendix: Proof of Theorem 1

From Equation (10) we note that

$$\begin{aligned} \tilde{b} \triangleq \hat{b} - b &= -\hat{Y}^{+}\hat{y} + S^{+}s \\ &= -\hat{Y}^{+}(s+\tilde{s}) + S^{+}s \\ &= -\left(\hat{Y}^{+} - S^{+}\right)s - \hat{Y}^{+}\tilde{s} \end{aligned} \tag{23}$$

where

$$\left[\begin{array}{c:c} \hat{y} & \hat{Y} \end{array}\right] = \left[\begin{array}{c:c} s & S \end{array}\right] + \left[\begin{array}{c:c} \tilde{s} & \tilde{S} \end{array}\right]. \tag{24}$$

Here, $\left[\begin{array}{c:c} s & S \end{array}\right]$ is the noise free version of $\left[\begin{array}{c:c} \hat{y} & \hat{Y} \end{array}\right]$, and b is the Lth order linear prediction coefficient vector associated with the poles of the noiseless model. Note that b gives the n true poles and $L-n$ extraneous poles (*cf.* [34]). In order to evaluate the expression in Equation (23) we use the following identity for any matrix M [35]

$$\widehat{M}^{+} - M^{+} = -\widehat{M}^{+}\widetilde{M}M^{+} + \left(\widehat{M}^{*}\widehat{M}\right)^{+}\widetilde{M}^{*}P_{M}^{\perp} + P_{\widehat{M}^{*}}^{\perp}\widetilde{M}^{*}\left(MM^{*}\right)^{+}, \tag{25}$$

where $\widetilde{M} = \widehat{M} - M$, $P_{M}^{\perp} = I_m - MM^{+}$, $P_{\widehat{M}^{*}}^{\perp} = I_n - \widehat{M}^{*}\widehat{M}^{*+}$, and I_q is the $q \times q$ identity matrix. Using the fact that $P_{S}^{\perp}s = 0$ and a first order approximation, we then obtain

$$\begin{aligned} \tilde{b} &= -\hat{Y}^{+}\tilde{S}b - P_{\hat{Y}^{*}}^{\perp}\tilde{S}^{*}S^{*+}S^{+}s - \hat{Y}^{+}\tilde{s} \\ &\approx -S^{+}\tilde{S}b - P_{S^{*}}^{\perp}\tilde{S}^{*}S^{*+}S^{+}s - S^{+}\tilde{s}. \end{aligned} \tag{26}$$

The above approximation is valid since the matrices $\left[\begin{array}{cc} \hat{y} & \hat{Y} \end{array}\right]$ and $\left[\begin{array}{cc} s & S \end{array}\right]$ have the same rank. Let

$$\left[\begin{array}{cc} y & Y \end{array}\right] = \left[\begin{array}{cc} s & S \end{array}\right] + \left[\begin{array}{cc} w & W \end{array}\right], \tag{27}$$

where $\left[\begin{array}{cc} w & W \end{array}\right]$ is defined as the noise-only part of $\left[\begin{array}{cc} y & Y \end{array}\right]$ (see Equations (2) and (7)). By using the perturbation analysis in [36] on the matrices $\left[\begin{array}{cc} s & S \end{array}\right]$, $\left[\begin{array}{cc} y & Y \end{array}\right]$, and $\left[\begin{array}{cc} \hat{y} & \hat{Y} \end{array}\right]$, it can be shown that to a first order approximation

$$S^{+}\left[\begin{array}{cc} \tilde{s} & \tilde{S} \end{array}\right] = S^{+}\left[\begin{array}{cc} w & W \end{array}\right]. \tag{28}$$

The above equation also implies that $\tilde{S}^{*}S^{*+} = W^{*}S^{*+}$. Thus, Equation (26) can be written (to a first order approximation) as

$$\tilde{b} = -S^{+}Wb - P_{S}^{\perp}W^{*}S^{*+}S^{+}s - S^{+}w. \tag{29}$$

From Equation (29), we note that $\tilde{b} \rightarrow 0$ as $\sigma \rightarrow 0$ since the elements of W and w are uncorrelated, zero mean, complex white Gaussian random variables. Therefore, the resulting L pole estimates are consistent as $\sigma \rightarrow 0$. Similarly it can be shown that the L sets of the amplitude coefficients are also consistent as $\sigma \rightarrow 0$. Note that the "true" amplitude coefficients of the extraneous modes are zero; thus it follows that choosing the n highest energy poles as the true poles is consistent as $\sigma \rightarrow 0$.

Note that $SP_{S}^{\perp} = 0$. Multiplying both sides of Equation (29) through by S, we obtain

$$S\tilde{b} = -SS^{+}\epsilon, \tag{30}$$

where $\epsilon = w + Wb$.

From Equation (2), we can write S as

$$S = \left[\begin{array}{cccc} x_1(1) & x_2(1) & \cdots & x_n(1) \\ x_1(1)p_1 & x_2(1)p_2 & \cdots & x_n(1)p_n \\ \vdots & \vdots & & \vdots \\ x_1(1)p_1^{m-(L+1)} & x_2(1)p_2^{m-(L+1)} & \cdots & x_n(1)p_n^{m-(L+1)} \\ \hline x_1(2) & x_2(2) & \cdots & x_n(2) \\ x_1(2)p_1 & x_2(2)p_2 & \cdots & x_n(2)p_n \\ \vdots & \vdots & & \vdots \\ x_1(2)p_1^{m-(L+1)} & x_2(2)p_2^{m-(L+1)} & \cdots & x_n(2)p_n^{m-(L+1)} \\ \hline & & \vdots & \\ \hline x_1(N) & x_2(N) & \cdots & x_n(N) \\ x_1(N)p_1 & x_2(N)p_2 & \cdots & x_n(N)p_n \\ \vdots & \vdots & & \vdots \\ x_1(N)p_1^{m-(L+1)} & x_2(N)p_2^{m-(L+1)} & \cdots & x_n(N)p_n^{m-(L+1)} \end{array}\right] G, \tag{31}$$

or

$$S = HG, \tag{32}$$

where G is given by

$$G = \begin{bmatrix} p_1 & p_1^2 & \cdots & p_1^L \\ p_2 & p_2^2 & \cdots & p_2^L \\ \vdots & \vdots & & \vdots \\ p_n & p_n^2 & \cdots & p_n^L \end{bmatrix}. \tag{33}$$

Equation (30) thus becomes

$$HG\tilde{b} = -HGS^{+}\epsilon. \tag{34}$$

Now note from Equation (8) that the true and estimated Lth order characteristic polynomials are $B(z) = 1+b_1z^1+b_2z^2+\cdots+b_Lz^L$ and $\widehat{B}(z) = 1+\widehat{b}_1z^1+\widehat{b}_2z^2+\cdots+\widehat{b}_Lz^L$,

respectively. Note that $B(p_i) = 0$ and $\widehat{B}(\widehat{p}_i) = 0$.

We can use a first order Taylor expansion to find an expression for the error in the estimated pole locations. We follow the technique in [17]. For each $\widehat{p}_i$ we obtain

$$
\begin{aligned}
0 &= \widehat{B}(\widehat{p}_i) \\
&= \widehat{B}(p_i) + \frac{\partial}{\partial z}\widehat{B}(z)|_{z=p_i}(\widehat{p}_i - p_i) + (\text{higher order terms}) \\
&= \widehat{B}(p_i) - B(p_i) + \frac{\partial}{\partial z}\widehat{B}(z)|_{z=p_i}(\widehat{p}_i - p_i) + (\text{higher order terms}) \\
&\approx 1 + \widehat{b}_1 p_i + \widehat{b}_2 p_i^2 + \cdots + \widehat{b}_L p_i^L - \left(1 + b_1 p_i + b_2 p_i^2 + \cdots + b_L p_i^L\right) \\
&\quad + \left(\widehat{b}_1 + 2\widehat{b}_2 p_i + \cdots + L\widehat{b}_L p_i^{L-1}\right)(\widehat{p}_i - p_i) \\
&\approx \begin{bmatrix} p_i & p_i^2 & \cdots & p_i^L \end{bmatrix} \begin{bmatrix} \widehat{b}_1 - b_1 \\ \widehat{b}_2 - b_2 \\ \vdots \\ \widehat{b}_L - b_L \end{bmatrix} + \begin{bmatrix} b_1 & b_2 & \cdots & b_L \end{bmatrix} \begin{bmatrix} 1 \\ 2p_i \\ \vdots \\ Lp_i^{L-1} \end{bmatrix} (\widehat{p}_i - p_i) \\
&= \begin{bmatrix} p_i & p_i^2 & \cdots & p_i^L \end{bmatrix} \left(\widehat{b} - b\right) + \eta_i (\widehat{p}_i - p_i). \qquad (35)
\end{aligned}
$$

Thus, to a first-order approximation

$$
(\widehat{p}_i - p_i) = -\frac{1}{\eta_i} \begin{bmatrix} p_i & p_i^2 & \cdots & p_i^L \end{bmatrix} \widetilde{b}, \qquad (36)
$$

where η_i is given by

$$
\eta_i = \begin{bmatrix} b_1 & b_2 & \cdots & b_L \end{bmatrix} \begin{bmatrix} 1 \\ 2p_i \\ \vdots \\ Lp_i^{L-1} \end{bmatrix}. \qquad (37)
$$

In matrix form we thus obtain

$$\tilde{P} \triangleq \hat{P} - P = -FG\tilde{b}, \tag{38}$$

where

$$\begin{aligned} \hat{P} &= \begin{bmatrix} \hat{p}_1 & \hat{p}_2 & \cdots & \hat{p}_n \end{bmatrix}^T \\ P &= \begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix}^T \\ F &= \operatorname{diag}\left(\frac{1}{\eta_1}, \frac{1}{\eta_2}, \ldots, \frac{1}{\eta_n}\right). \end{aligned} \tag{39}$$

Since the n poles are distinct, H is of full column rank. Hence, we can multiply Equation (34) by $(H^*H)^{-1}H^*$ to get

$$G\tilde{b} = -GS^+\epsilon, \tag{40}$$

and by substituting Equation (40) into Equation (38) we obtain

$$\tilde{P} = FGS^+\epsilon. \tag{41}$$

We now note that to a first order approximation, $\tilde{P}$ is given by

$$\begin{aligned} \tilde{P} &= \hat{\alpha} \odot e^{j\hat{\omega}} - \alpha \odot e^{j\omega} \\ &= (\alpha + \tilde{\alpha}) \odot e^{j(\omega+\tilde{\omega})} - \alpha \odot e^{j\omega} \\ &= (\alpha + \tilde{\alpha}) \odot e^{j\omega} \odot (1 + j\tilde{\omega} + (\text{h.o.t.})) - \alpha \odot e^{j\omega} \\ &\approx (\alpha + \tilde{\alpha}) \odot e^{j\omega} + \alpha \odot e^{j\omega} \odot j\tilde{\omega} - \alpha \odot e^{j\omega} \\ &= T_p^{-1}(T_\alpha \tilde{\alpha} + j\tilde{\omega}), \end{aligned} \tag{42}$$

where $\odot$ denotes the Hadamard product, and

$$\begin{aligned} e^{j\omega} &= \left[\begin{array}{cccc} e^{j\omega_1} & e^{j\omega_2} & \cdots & e^{j\omega_n} \end{array} \right]^T \\ \tilde{\omega} &= \left[\begin{array}{cccc} \tilde{\omega}_1 & \tilde{\omega}_2 & \cdots & \tilde{\omega}_n \end{array} \right]^T \\ \tilde{\alpha} &= \left[\begin{array}{cccc} \tilde{\alpha}_1 & \tilde{\alpha}_2 & \cdots & \tilde{\alpha}_n \end{array} \right]^T \\ T_p &= \operatorname{diag}\left(\frac{1}{p_1}, \frac{1}{p_2}, \ldots, \frac{1}{p_n}\right) \end{aligned} \tag{43}$$

and where $\tilde{\omega}_i = \hat{\omega}_i - \omega_i$ and $\tilde{\alpha}_i = \hat{\alpha}_i - \alpha_i$. From Equations (41) and (42) we obtain:

$$\begin{aligned} \tilde{\omega} &= \operatorname{Im}\left\{T_p F G S^+ \epsilon\right\} \\ \tilde{\alpha} &= \operatorname{Re}\left\{T_\alpha^{-1} T_p F G S^+ \epsilon\right\}. \end{aligned} \tag{44}$$

Recall that the elements of W and w are uncorrelated, zero mean, complex white Gaussian random variables. Thus, ϵ is multivariate Gaussian with zero mean and covariance matrix

$$E\left[\epsilon\epsilon^*\right] = E\left[\left(\left[\begin{array}{ccc} w & : & W \end{array}\right]\left[\begin{array}{c} 1 \\ b \end{array}\right]\right)\left(\left[\begin{array}{ccc} w & : & W \end{array}\right]\left[\begin{array}{c} 1 \\ b \end{array}\right]\right)^*\right] = \sigma D D^*. \tag{45}$$

D is defined as a $(m-L)N \times mN$ block diagonal matrix given by

$$D = \operatorname{diag}\left(B, B, \ldots, B\right), \tag{46}$$

where B is given by

$$B = \begin{bmatrix} 1 & b_1 & b_2 & \cdots & b_L & 0 & 0 & \cdots & 0 \\ 0 & 1 & b_1 & \cdots & b_{L-1} & b_L & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & b_1 & \cdots & b_{L-1} & b_L & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & b_{L-2} & b_{L-1} & b_L \end{bmatrix}_{(m-L)\times(m)} . \quad (47)$$

We also have

$$E\left[\epsilon\epsilon^T\right] = E\left[\left(\begin{bmatrix} w & : & W \end{bmatrix}\begin{bmatrix} 1 \\ b \end{bmatrix}\right)\left(\begin{bmatrix} w & : & W \end{bmatrix}\begin{bmatrix} 1 \\ b \end{bmatrix}\right)^T\right] = 0. \quad (48)$$

Using these results, along with the following relationships (proven in [16])

$$\begin{aligned} \mathrm{Re}\{u\}\mathrm{Re}\{v^T\} &= \frac{1}{2}\left[\mathrm{Re}\{uv^T\} + \mathrm{Re}\{uv^*\}\right] \\ \mathrm{Im}\{u\}\mathrm{Im}\{v^T\} &= -\frac{1}{2}\left[\mathrm{Re}\{uv^T\} - \mathrm{Re}\{uv^*\}\right] \\ \mathrm{Re}\{u\}\mathrm{Im}\{v^T\} &= \frac{1}{2}\left[\mathrm{Im}\{uv^T\} - \mathrm{Im}\{uv^*\}\right], \end{aligned} \quad (49)$$

we obtain the following covariances for the pole parameters:

$$\begin{aligned} E\left[\tilde{\omega}\tilde{\omega}^T\right] &= \frac{\sigma}{2}\mathrm{Re}\left\{T_p F G S^+ D D^* S^{*+} G^* F^* T_p^*\right\} \\ E\left[\tilde{\omega}\tilde{\alpha}^T\right] &= \frac{\sigma}{2}\mathrm{Im}\left\{T_p F G S^+ D D^* S^{*+} G^* F^* T_p^* T_\alpha^{*-1}\right\} \\ E\left[\tilde{\alpha}\tilde{\alpha}^T\right] &= \frac{\sigma}{2}\mathrm{Re}\left\{T_\alpha^{-1} T_p F G S^+ D D^* S^{*+} G^* F^* T_p^* T_\alpha^{*-1}\right\}. \end{aligned} \quad (50)$$

To obtain the covariances for the amplitude coefficient parameters we use Equation (14), which provides the amplitude coefficient estimates for each snapshot in terms of the estimated poles. We now note the following

$$\widetilde{X} \triangleq \widehat{X} - X = \hat{A}^+ Y_a - A^+ S_a$$

$$
\begin{aligned}
&= \hat{A}^{+}\left(S_a + N_a\right) - A^{+}S_a \\
&= \left(\hat{A}^{+} - A^{+}\right) S_a + \hat{A}^{+}N_a, \qquad (51)
\end{aligned}
$$

where

$$
X = \left[\begin{array}{cccc} x(1) & x(2) & \cdots & x(N) \end{array}\right], \qquad (52)
$$

and where S_a is the noise free version of Y_a and $N_a = \left[\begin{array}{cccc} e(1) & e(2) & \cdots & e(N) \end{array}\right]$ is the corresponding noise.

We apply the identity in Equation (25) to the first term of Equation (51). Since the n poles are distinct, A and $\hat{A}$ are full rank. Also, since $m > n$ and $S_a \in \text{Range}(A)$, we have

$$
(\hat{A}^{+} - A^{+})S_a = -\hat{A}^{+}\tilde{A}A^{+}S_a. \qquad (53)
$$

From Equations (51) and (53) we obtain the following first order approximation

$$
\begin{aligned}
\widetilde{X} &= -\hat{A}^{+}\tilde{A}A^{+}S_a + \hat{A}^{+}N_a \\
&= -\hat{A}^{+}\tilde{A}X + \hat{A}^{+}N_a \\
&\approx -A^{+}\tilde{A}X + A^{+}N_a. \qquad (54)
\end{aligned}
$$

Note that $\widetilde{X}$ is a matrix in which each column is composed of the amplitude coefficient variations for each snapshot:

$$
\widetilde{X} = \left[\begin{array}{cccc} \tilde{x}(1) & \tilde{x}(2) & \cdots & \tilde{x}(N) \end{array}\right]. \qquad (55)
$$

Following the same procedure as in Equation (42), a first order approximation of $\tilde{x}(t)$ is given by

$$
\tilde{x}(t) \approx T_x^{-1}(t)\left(T_\beta(t)\tilde{\beta}(t) + j\tilde{\gamma}(t)\right), \qquad (56)
$$

where

$$\begin{aligned}
\tilde{\gamma}(t) &= \left[\begin{array}{cccc} \tilde{\gamma}_1(t) & \tilde{\gamma}_2(t) & \cdots & \tilde{\gamma}_n(t) \end{array}\right]^T \\
\tilde{\beta}(t) &= \left[\begin{array}{cccc} \tilde{\beta}_1(t) & \tilde{\beta}_2(t) & \cdots & \tilde{\beta}_n(t) \end{array}\right]^T \\
T_x(t) &= \operatorname{diag}\left(\frac{1}{x_1(t)}, \frac{1}{x_2(t)}, \ldots, \frac{1}{x_n(t)}\right)
\end{aligned} \tag{57}$$

and where $\tilde{\gamma}_i(t) = \hat{\gamma}_i(t) - \gamma_i(t)$ and $\tilde{\beta}_i(t) = \hat{\beta}_i(t) - \beta_i(t)$. From Equations (54), (55), and (56) we obtain

$$\begin{aligned}
\tilde{\gamma}(t) &= \operatorname{Im}\left\{T_x(t)A^+\left(-\tilde{A}x(t) + e(t)\right)\right\} \\
\tilde{\beta}(t) &= \operatorname{Re}\left\{T_\beta^{-1}(t)T_x(t)A^+\left(-\tilde{A}x(t) + e(t)\right)\right\}.
\end{aligned} \tag{58}$$

Before computing covariances for the parameters in Equation (58), we need to perform some manipulations to $\tilde{A}x(t)$, since the random variable $\tilde{A}$ does not appear at the rightmost position. We proceed with a first order approximation as follows

$$\begin{aligned}
\tilde{A}x(t) &= \left[\begin{array}{cccc}
0 & 0 & \cdots & 0 \\
\hat{p}_1 - p_1 & \hat{p}_2 - p_2 & \cdots & \hat{p}_n - p_n \\
\hat{p}_1^2 - p_1^2 & \hat{p}_2^2 - p_2^2 & \cdots & \hat{p}_n^2 - p_n^2 \\
\vdots & \vdots & & \vdots \\
\hat{p}_1^{m-1} - p_1^{m-1} & \hat{p}_2^{m-1} - p_2^{m-1} & \cdots & \hat{p}_n^{m-1} - p_n^{m-1}
\end{array}\right] x(t) \\
&\approx \left[\begin{array}{cccc}
0 & 0 & \cdots & 0 \\
\tilde{p}_1 & \tilde{p}_2 & \cdots & \tilde{p}_n \\
2p_1\tilde{p}_1 & 2p_2\tilde{p}_2 & \cdots & 2p_n\tilde{p}_n \\
\vdots & \vdots & & \vdots \\
(m-1)\,p_1^{m-2}\tilde{p}_1 & (m-1)\,p_2^{m-2}\tilde{p}_2 & \cdots & (m-1)\,p_n^{m-2}\tilde{p}_n
\end{array}\right] x(t)
\end{aligned}$$

$$
\begin{aligned}
&= CAT_p \mathrm{diag}\left(\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_n\right) x(t) \\
&= CAT_x^{-1}(t) T_p \tilde{P} \\
&= CAT_x^{-1}(t) T_p F G S^+ \epsilon, \qquad (59)
\end{aligned}
$$

where C is a diagonal matrix given by

$$
C = \mathrm{diag}\left(0, 1, \ldots, m-1\right). \qquad (60)
$$

Equation (58) can now be approximated by

$$
\begin{aligned}
\tilde{\gamma}(t) &\approx \mathrm{Im}\left\{T_x(t) A^+ \left(-CAT_x^{-1}(t) T_p F G S^+ \epsilon + e(t)\right)\right\} \\
\tilde{\beta}(t) &\approx \mathrm{Re}\left\{T_\beta^{-1}(t) T_x(t) A^+ \left(-CAT_x^{-1}(t) T_p F G S^+ \epsilon + e(t)\right)\right\}, \qquad (61)
\end{aligned}
$$

Since

$$
E\left[\epsilon e^*(t)\right] = E\left[\left(\begin{bmatrix} w & W \end{bmatrix} \begin{bmatrix} 1 \\ b \end{bmatrix}\right) e^*(t)\right] = \sigma D(t), \qquad (62)
$$

where $D(t)$ are each given by the tth column block of D (*cf.* see Equation (46)), we also have

$$
\begin{aligned}
E\left[\epsilon \epsilon^T(t)\right] &= E\left[\left(\begin{bmatrix} w & W \end{bmatrix} \begin{bmatrix} 1 \\ b \end{bmatrix}\right) e^T(t)\right] = 0 \\
E\left[e(t) e^T(t)\right] &= 0. \qquad (63)
\end{aligned}
$$

Using these results and Equations (49) and (61) we obtain the following covariances for the amplitude coefficient parameters

$$
\begin{aligned}
E\left[\tilde{\gamma}(t)\tilde{\gamma}^T(r)\right] = \frac{\sigma}{2} \mathrm{Re}\Big\{ & T_x(t) A^+ \Big(CAT_x^{-1}(t) T_p F G S^+ D D^* S^{+*} G^* F^* T_p^* T_x^{*-1}(r) A^* C^* \\
& - CAT_p T_x^{-1}(t) F G S^+ D(r) - D^*(t) S^{+*} G^* F^* T_p^* T_x^{*-1}(r) A^* C^* \\
& + I_m \delta_{t,r}\Big) A^{+*} T_x^*(r)\Big\}
\end{aligned}
$$

$$
\begin{aligned}
E\left[\tilde{\gamma}(t)\tilde{\beta}^T(r)\right] &= \frac{\sigma}{2}\mathrm{Im}\Big\{T_x(t)A^+\Big(CAT_x^{-1}(t)T_pFGS^+DD^*S^{+*}G^*F^*T_p^*T_x^{*-1}(r)A^*C^* \\
&\quad -CAT_x^{-1}(t)T_pFGS^+D(r) - D^*(t)S^{+*}G^*F^*T_p^*T_x^{*-1}(r)A^*C^* \\
&\quad +I_m\delta_{t,r}\Big)A^{+*}T_x^*(r)T_\beta^{-1*}(r)\Big\} \\
E\left[\tilde{\beta}(t)\tilde{\beta}^T(r)\right] &= \frac{\sigma}{2}\mathrm{Re}\Big\{T_\beta^{-1}(t)T_x(t)A^+\Big(CAT_x^{-1}(t)T_pFGS^+DD^*S^{+*}G^*F^*T_p^*T_x^{*-1}(r)A^*C^* \\
&\quad -CAT_x^{-1}(t)T_pFGS^+D(r) - D^*(t)S^{+*}G^*F^*T_p^*T_x^{*-1}(r)A^*C^* \\
&\quad +I_m\delta_{t,r}\Big)A^{+*}T_x^*(r)T_\beta^{-1*}(r)\Big\}. \qquad (64)
\end{aligned}
$$

Using Equations (44), (49), and (61) we can also compute the cross-covariances between the poles and the amplitude coefficients as follows

$$
\begin{aligned}
E\left[\tilde{\gamma}(t)\tilde{\omega}^T\right] &= \frac{\sigma}{2}\mathrm{Re}\Big\{T_x(t)A^+\Big(-CAT_x^{-1}(t)T_pFGS^+DD^* + D^*(t)\Big)S^{+*}G^*F^*T_p^*\Big\} \\
E\left[\tilde{\gamma}(t)\tilde{\alpha}^T\right] &= \frac{\sigma}{2}\mathrm{Im}\Big\{T_x(t)A^+\Big(-CAT_x^{-1}(t)T_pFGS^+DD^* + D^*(t)\Big)S^{+*}G^*F^*T_p^*T_\alpha^{-1*}\Big\} \\
E\left[\tilde{\beta}(t)\tilde{\omega}^T\right] &= -\frac{\sigma}{2}\mathrm{Im}\Big\{T_\beta^{-1}(t)T_x(t)A^+\Big(-CAT_x^{-1}(t)T_pFGS^+DD^* + D^*(t)\Big)S^{+*}G^*F^*T_p^*\Big\} \\
E\left[\tilde{\beta}(t)\tilde{\alpha}^T\right] &= \frac{\sigma}{2}\mathrm{Re}\Big\{T_\beta^{-1}(t)T_x(t)A^+\Big(-CAT_x^{-1}(t)T_pFGS^+DD^* + D^*(t)\Big)S^{+*}G^*F^*T_p^*T_\alpha^{-1*}\Big\}. \\
&\qquad (65)
\end{aligned}
$$

Equations (50), (64), and (65) completely specify Σ_θ as given in Equation (20) using the following substitutions

$$
\begin{aligned}
U(t,r) &= R(t)ZR^*(r) - R(t)Q(r) - Q^*(t)R^*(r) + \frac{\sigma}{2}T_x(t)A^+A^{+*}T_x^*(r)\delta_{t,r} \\
V(t) &= -R(t)Z + Q^*(t) \\
Z &= \frac{\sigma}{2}T_pFGS^+DD^*S^{*+}G^*F^*T_p^* \\
Q(t) &= \frac{\sigma}{2}T_pFGS^+D(t)A^{+*}T_x^*(t) \\
R(t) &= T_x(t)A^+CAT_x^{-1}(t). \qquad (66)
\end{aligned}
$$

Furthermore, since ϵ and $e(t)$ are zero mean, Equations (44) and (61) imply that the parameter estimates are unbiased, *i.e.* to a first order approximation $E\left[\widehat{\theta}\right] = \theta$. This completes the proof. □

References

[1] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. RADC Spectral Estimation Workshop*, (Rome, NY), pp. 243–258, 1979.

[2] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM*, vol. 17, no. 6, pp. 883–893, December 1980.

[3] T. L. Henderson, "Geometric methods for determining system poles from transient response," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 5, pp. 982–988, October 1981.

[4] R. Kumaresan and D. W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, no. 6, pp. 833–840, December 1982.

[5] R. Roy, A. Paulraj, and T. Kailath, "ESPRIT–a subspace rotation approach to estimation of parameters of cisoids in noise," *IEEE Transactions on Acoustics, Speech. and Signal Processing*, vol. ASSP-34, no. 4, pp. 1340–1342, April 1986.

[6] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–34, no. 5, pp. 1361–1375, October 1986.

[7] R. Kumaresan, L. L. Scharf, and A. K. Shaw, "An algortihm for pole-zero modeling and analysis," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–34, no. 6, pp. 637–640, June 1986.

[8] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–34, no. 2, pp. 331–341, April 1986.

[9] R. Kumaresan and D. Tufts, "Accurate parameter estimation of noisy speech-like signals," in *Proc. Internat. Conf. Acoust. Speech. Signal Process. '82*, vol. 3, pp. 1357–1361, 1982.

[10] M. A. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–35, no. 10, pp. 1440–1454, October 1987.

[11] B. D. Rao, "Perturbation analysis of an SVD-based linear prediction method for estimating the frequencies of multiple sinusoids," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-36, no. 7, pp. 1026–1035, July 1988.

[12] Y. Hua and T. K. Sarkar, "Perturbation analysis of TK method for harmonic retrieval problems," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-36, no. 2, pp. 228–240, February 1988.

[13] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–36, no. 10, pp. 1553–1560, October 1988.

[14] R. Roy and T. Kailath, "ESPRIT– estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 7, pp. 984–995, July 1989.

[15] M. D. Zoltowski and D. Stavrinides, "Sensor array signal processing via a Procrustes rotations based eigenanalysis of the ESPRIT data pencil," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 6, pp. 832–861, June 1989.

[16] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramér-Rao Bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–37, no. 5, pp. 720–741, May 1989.

[17] P. Stoica, T. Söderström, and F. Ti, "Asymptotic properties of the high-order Yule-Walker estimates of sinusoidal frequencies," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–37, no. 11, pp. 1721–1734, November 1989.

[18] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, no. 5, pp. 814–824, May 1990.

[19] E. M. Dowling and R. D. DeGroat, "The equivalence of the total least squares and minimum norm methods," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-39, no. 8, pp. 1891–1892, August 1991.

[20] B. Ottersten, M. Viberg, and T. Kailath, "Performance analysis of the total least squares ESPRIT algorithm," *IEEE Transactions on Signal Processing*, vol. SP-39, no. 5, pp. 1122–1135, May 1991.

[21] P. Stoica and A. Nehorai, "Performance comparison of subspace rotation and MUSIC methods for direction estimation," *IEEE Transactions on Signal Processing*, vol. SP-31, no. 2, pp. 446–453, February 1991.

[22] T. J. Abatzoglou and L. K. Lam, "Direction finding using uniform arrays and the constrained total least squares method," in *Proceedings of the Twenty-Fifth Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA), November 4–6 1991.

[23] H. Clergeot, S. Tressens, and A. Ouamri, "Performance of high resolution frequencies estimation methods compared to the Cramér-rao bounds," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 11, pp. 1703–1720, November 1989.

[24] R. Vaccaro, D. Tufts, and G. Boudreaux-Bartels, "Advances in principal component signal processing," in *SVD and Signal Processing — Algorithms, Applications and architectures (edited by E. Deprettere)*, pp. 115–146, Elsevier Science Publishers B.V.(North-Holland), 1988.

[25] A. Kot, S. Parthasarathy, D. Tufts, and R. Vaccaro, "Statistical performance of single sinusoid frequency estimation in white noise using state-variable balancing and linear prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 11, pp. 1639–1642, November 1987.

[26] F. Li and R. Vaccaro, "Analysis of min-norm and MUSIC with arbitrary array geometry," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-38, no. 6, pp. 976–985, November 1990.

[27] D. Tufts, A. Kot, and R. Vaccaro, "The threshold effect in signal processing algorithms which use an estimated subspace," in *SVD and Signal Processing, II-Algorithms, Analysis and Applications (edited by R. Vaccaro)*, pp. 301–320, Elsevier Science Publishers B.V.(North-Holland), 1991.

[28] B. Porat and B. Friedlander, "On the accuracy of the Kumaresan-Tufts method for estimating complex damped exponentials," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 2, pp. 231–235, February 1987.

[29] P. Stoica and T. Söderdtröm, "On spectral and root forms of sinusoidal frequency estimators," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, (Toronto, Ontario), pp. 3257–3260, May 14–17, 1991.

[30] J. Sacchini, W. M. Steedly, and R. L. Moses, "Two-dimensional Prony modeling and parameter estimation," in *Proceedings of the International Conference on*

Acoustics, Speech, and Signal Processing, (San Francisco), pp. III: 333–336, April 1992.

[31] J. J. Sacchini, *Development of Two-Dimensional Parametric Radar Signal Modeling and Estimation Techniques with Application to Target Identification.* PhD thesis, The Ohio State University, March 1992.

[32] W. M. Steedly and R. L. Moses, "The Cramér-Rao bound for pole and amplitude estimates of damped exponential signals in noise," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, (Toronto, Ontario), pp. 3569–3572, May 14–17, 1991.

[33] W. M. Steedly, C. J. Ying, and R. L. Moses, "A modified TLS-Prony method using data decimation," in *Proceedings of the SPIE OE/AEROSPACE Science and Sensing Conference*, (Orlando, FL), April 22–24, 1992.

[34] R. Kumaresan, "On the zeros of the linear prediction-error filter for deterministic signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-31, no. 1, pp. 217–220, February 1983.

[35] P.-A. Wedin, "Perturbation theory for pseudo-inverses," *BIT*, vol. 13, pp. 217–232, 1973.

[36] Y. Hua and T. K. Sarkar, "A perturbation property of the TLS-LP method," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, no. 11, pp. 2004–2005, November 1990.

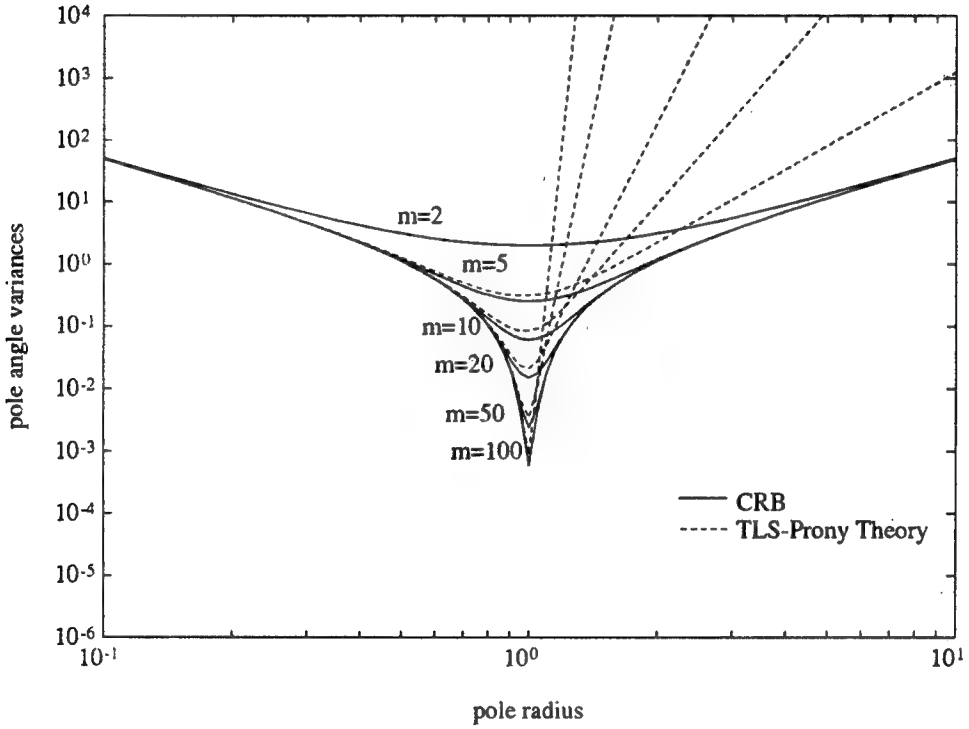


Figure 1: Pole angle variances for single pole data ($n = 1$).

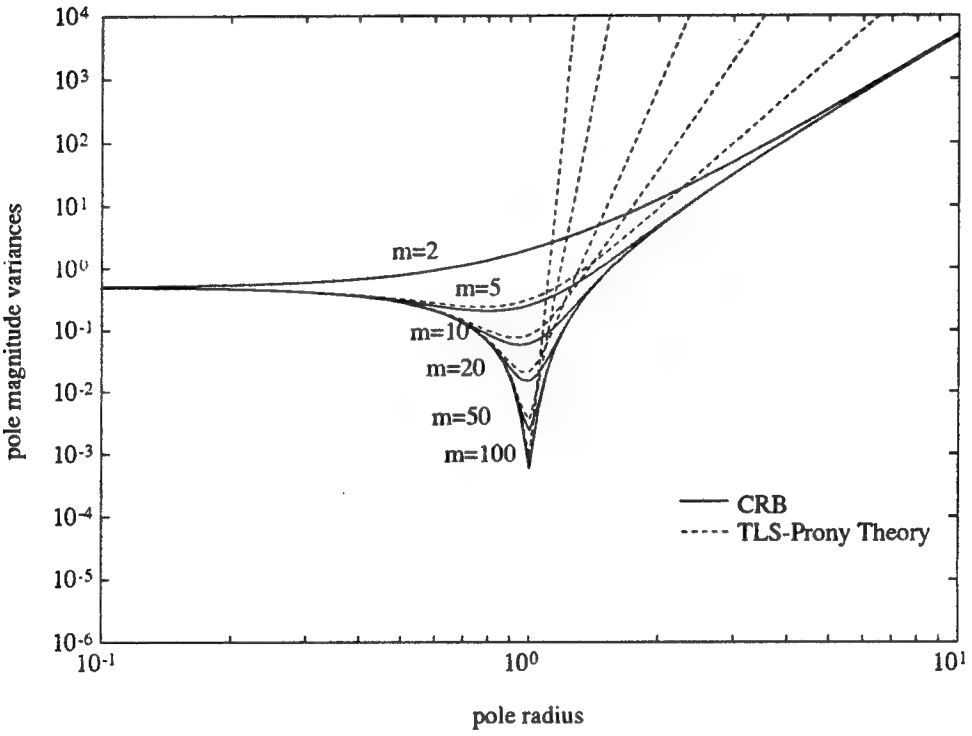


Figure 2: Pole magnitude variances for single pole data ($n = 1$).

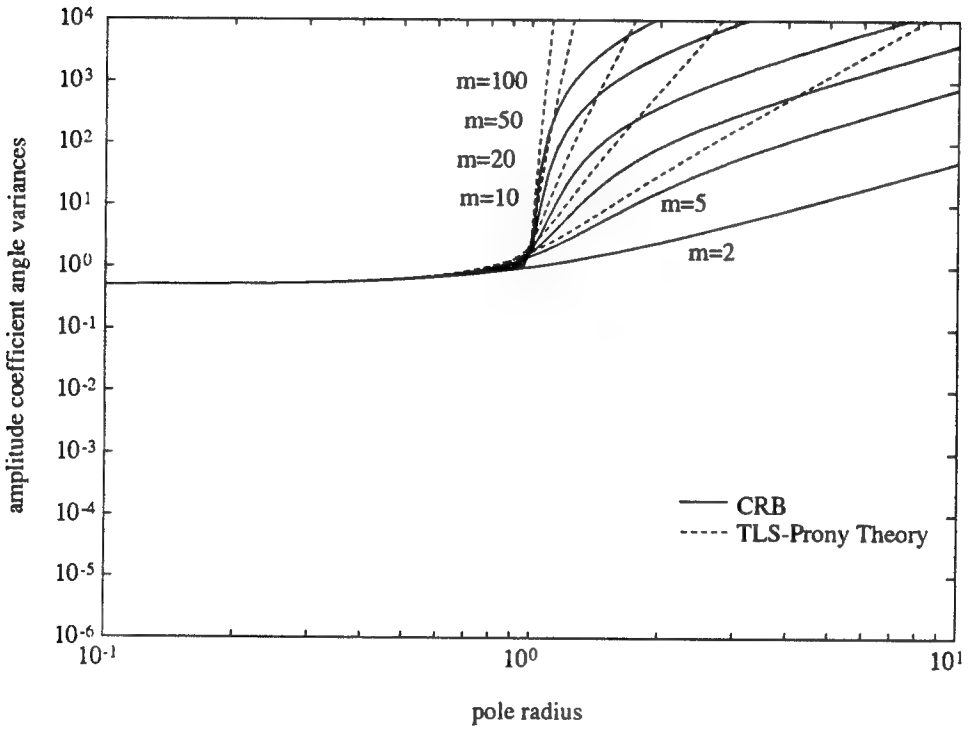


Figure 3: Amplitude coefficient angle variances for single pole data (n=1).

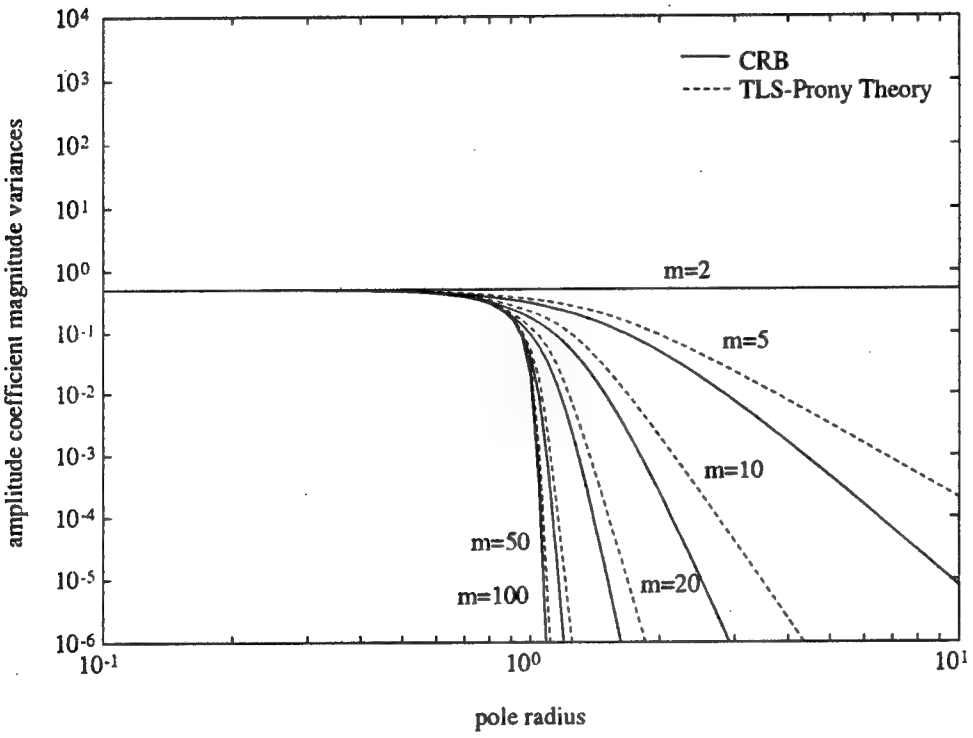


Figure 4: Amplitude coefficient magnitude variances for single pole data (n=1).

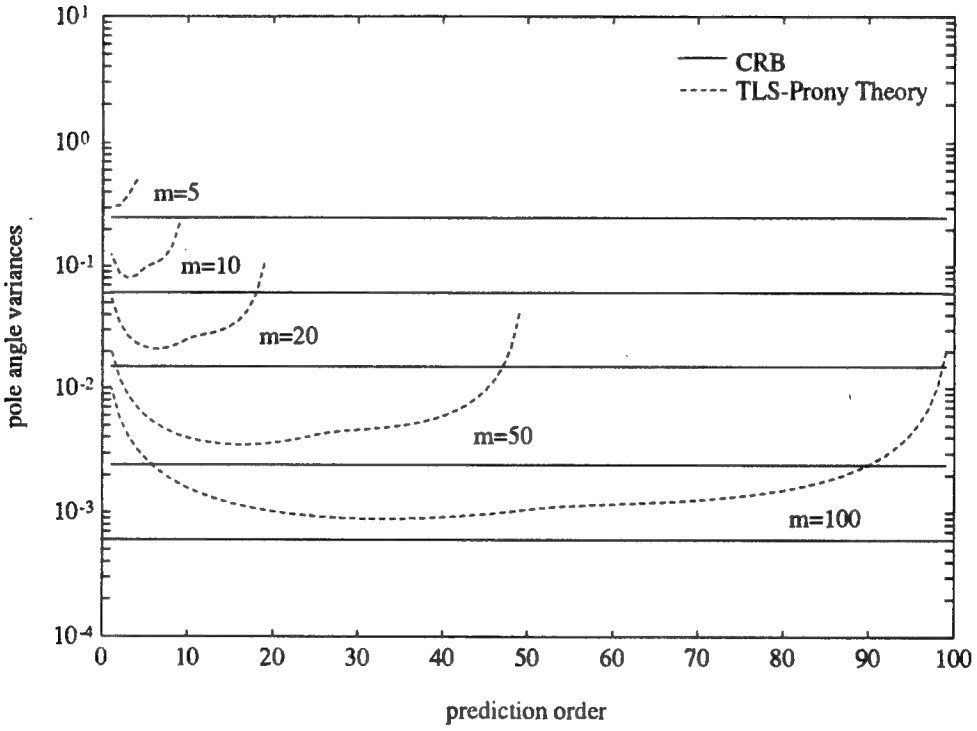


Figure 5: Pole angle variances for various prediction orders.

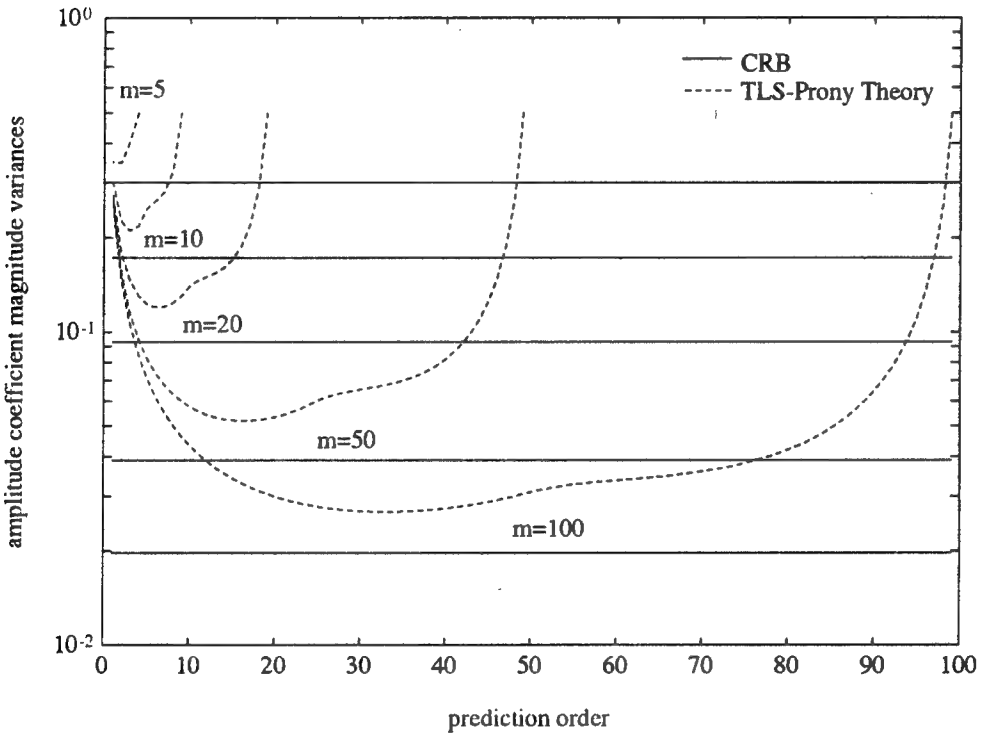


Figure 6: Amplitude coefficient magnitude variances for various prediction orders.

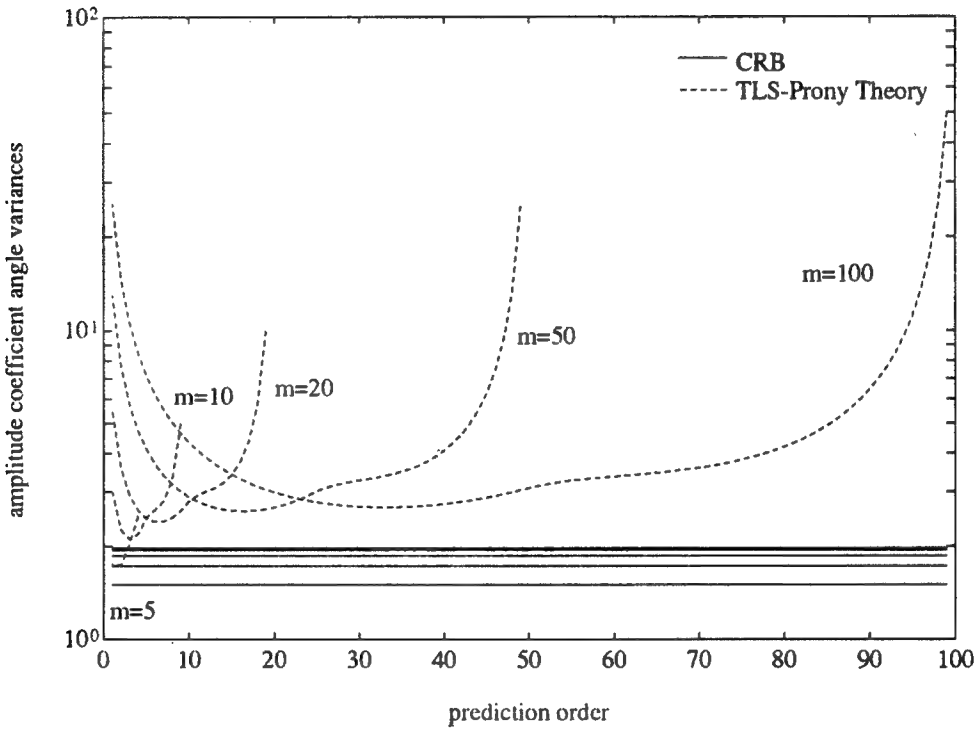


Figure 7: Amplitude coefficient angle variances for various prediction orders.

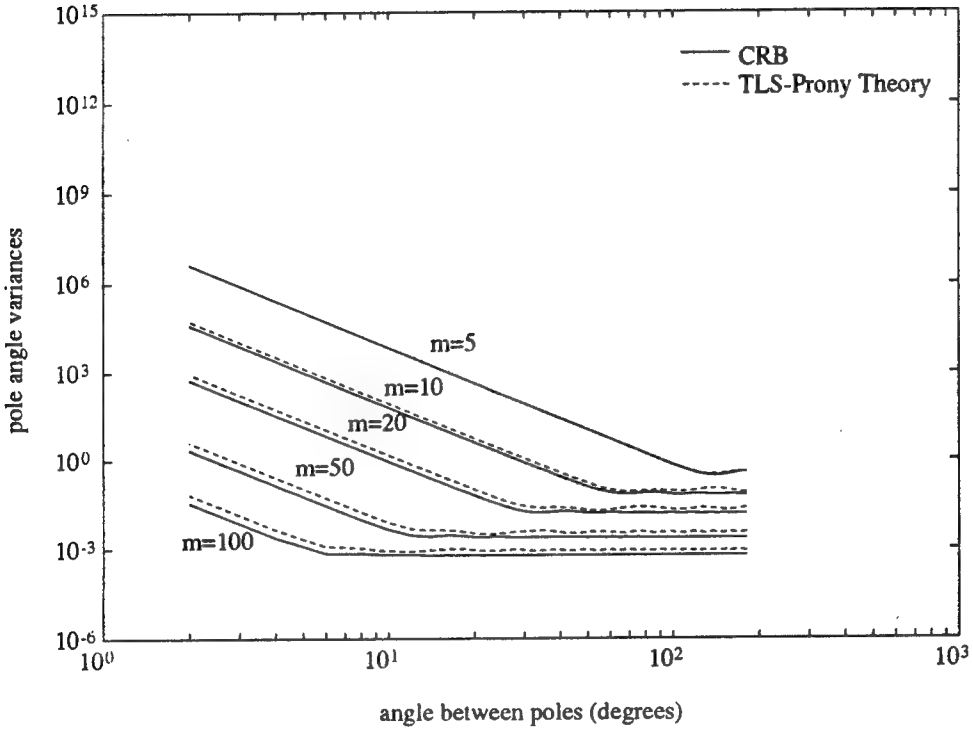
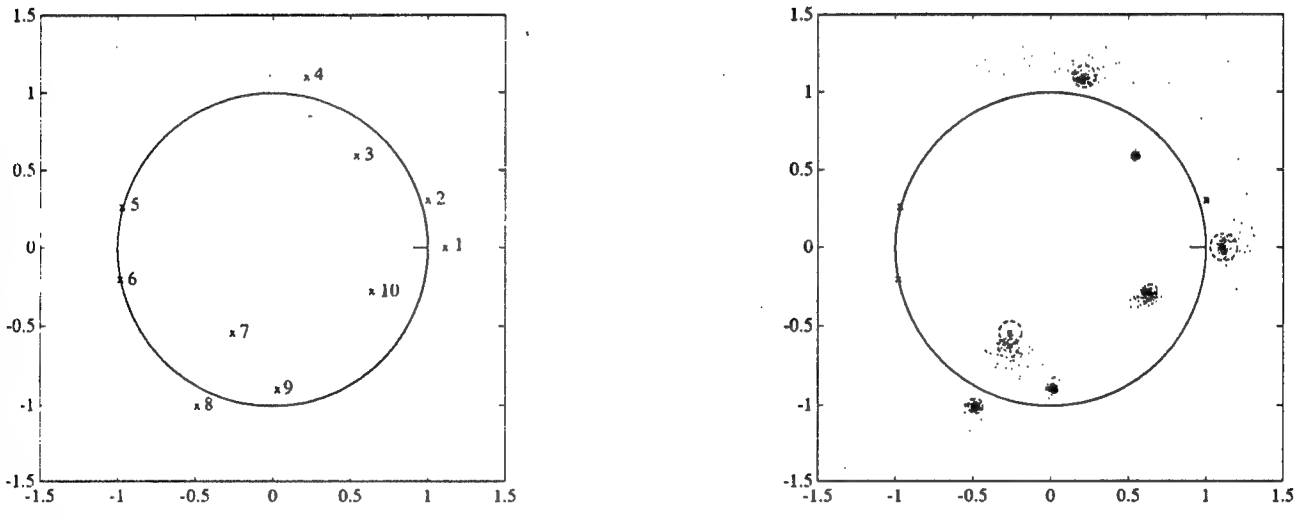


Figure 8: Angle variances as a function of pole separation for two poles on the unit circle.



(a) (b)

Figure 9: A general ten mode case: (a) true pole locations and (b) two standard deviation theoretical bounding circles and Monte-Carlo estimates.

B. Preprint of "A Modified TLS-Prony Method Using Data Decimation for Computationally Efficient Estimation of Damped Exponentials in Noise"

The following pages contain a preprint of the paper, "A Modified TLS-Prony Method Using Data Decimation for Computationally Efficient Estimation of Damped Exponentials in Noise," which has been submitted to *IEEE Transactions on Signal Processing*.

A Modified TLS-Prony Method Using Data Decimation †

William M. Steedly
The Analytic Sciences Corporation
Reston, VA 22090

Ching-Hui J. Ying and Randolph L. Moses
Department of Electrical Engineering
The Ohio State University
Columbus, OH 43210

Abstract

This paper introduces a modified TLS-Prony method which incorporates data decimation. The use of data decimation results in the reduction in the computational complexity because one high order estimation is replaced by several low order estimations. We present an analysis of pole variance statistics for this modified TLS-Prony method. This analysis provides a quantitative comparison of the parameter estimation accuracy as a function of decimation factors. We show that by using decimation, one can obtain comparable statistical performance results at a fraction of the computational cost, when compared to the conventional TLS-Prony algorithm.

Submitted to
IEEE Transactions on Signal Processing
August 1991

Revised January 1993

Version 1.1

EDICS Category 5.2.3, 5.1.5

Permission to publish this abstract separately is granted.

†This research was supported in part by the Air Force Office of Scientific Research, Bolling AFB, DC, and in part by the Avionics Division, Wright Laboratories, Wright Patterson AFB, OH.

I. Introduction

A popular high resolution estimation technique is the use of backward linear prediction coupled with singular value decomposition (SVD) and total least squares [1], here called the TLS-Prony technique. This technique has been shown to provide good parameter estimates of damped exponential signals in noise for various types of data [1, 2]. However, for large data lengths, the TLS-Prony method can be computationally expensive. The reason for this is that the TLS-Prony method involves computing the Singular Value Decomposition (SVD) of a data matrix of size (m, n) where m is related to data length and n to prediction order. For best accuracy of the parameter estimates, $n \approx \frac{m}{3}$. Thus, this data matrix becomes quite large for large data lengths. To overcome this problem it is sometimes possible to decimate the data before applying the TLS-Prony technique; the result is often a large reduction in computations. In this paper we consider the statistical and computational properties of the TLS-Prony algorithm when used in conjunction with data decimation.

Data decimation has been considered before [3, 4] in the context of spectral estimation. This technique entails using only part of the measured data. Decimation of correlation sequences was also considered in [5]; this technique effectively uses all the measured data, but is somewhat restrictive in that it applies only to correlation-based parameter estimation techniques. These works do not present a quantitative analysis of statistical properties of the resulting parameter estimates.

In this paper we develop a data decimation technique which is based on the TLS-

Prony algorithm [1]. We also present a theoretical statistical analysis of the accuracy of the TLS-Prony parameter estimates when decimation and any linear FIR filter are used. Based on this analysis, we present a quantitative comparison of estimation accuracy for various types of data decimation schemes. In particular, we compare the decimated and nondecimated procedures in terms of estimation accuracy. Our analysis demonstrates that the performance of the decimated TLS-Prony procedure is comparable to the performance of the nondecimated TLS-Prony procedure for undamped exponential modes. We also develop a complexity analysis and show that the decimated algorithm is computationally more efficient than the nondecimated algorithm. Both the statistical performance and the decrease in computational complexity are verified by Monte-Carlo simulations.

We then apply the statistical analysis to consider two specific cases of interest. First, the signals of interest may be bandlimited and occupy a relatively small region of the unambiguous frequency range $f \in \left[-\frac{1}{2}, \frac{1}{2}\right]$. In this case one is interested in analyzing a subset of the whole frequency range. For example, this technique was used to investigate radar signatures of aircraft [2, 6]. One question is whether decimation can improve the accuracy of the pole estimates in this case.

A second case of interest is when the signal occupies most or all of the unambiguous frequency range. In this case we filter the data to isolate a number of subbands, then use a decimation version of TLS-Prony to estimate the poles in each of the subbands. This idea is similar in principle to beamspace prefiltering in array processing [7, 8, 9]. By focusing on particular bands one at a time, estimation techniques can be used

with lower model orders since there are typically fewer modes within each of the bands. Thus, a single wideband estimation procedure is replaced by several lower order estimations. This has the advantages of being much less numerically intensive and of being amenable to parallel implementation [9].

An outline of this paper is as follows. In Section II we develop the modified TLS-Prony procedure. In Section III we derive the first order approximation of the statistics of the estimated parameters. In Section IV we develop a procedure and a complexity analysis for performing full spectrum estimation. We also discuss filter design and performance loss in the estimation. Section V presents some simulation studies using decimation. Finally, Section VI concludes the paper.

II. Decimation Estimation Procedure

A. Data Model

Assume we have N "snapshots" of data vectors $y(t)$, each of length m:

$$y(t) = \left[\begin{array}{cccc} y_0(t) & y_1(t) & \cdots & y_{m-1}(t) \end{array} \right]^T \qquad t = 1, 2, \ldots, N. \tag{1}$$

Each data vector is modeled as a noisy exponential sequence

$$y_q(t) = \sum_{i=1}^{n} x_i(t) p_i^q + e_q(t) \qquad q = 0, 1, \ldots, m-1. \tag{2}$$

There are n exponential modes in the data; the n poles $\{p_i\}_{i=1}^n$ do not vary from snapshot to snapshot, but the amplitudes $x_i(t)$ may vary. Here, it is assumed that

$\{e_q(t)\}$ are uncorrelated zero mean complex white Gaussian noise sequences with variance σ. Equation (2) may be compactly written as

$$y(t) = Ax(t) + e(t), \tag{3}$$

where $e(t) = \begin{bmatrix} e_0(t) & e_1(t) & \cdots & e_{m-1}(t) \end{bmatrix}^T$, $x(t) = \begin{bmatrix} x_0(t) & x_1(t) & \cdots & x_{n-1}(t) \end{bmatrix}^T$, and A is the $m \times n$ Vandermonde matrix derived from n signal poles

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ p_1 & p_2 & \cdots & p_n \\ p_1^2 & p_2^2 & \cdots & p_n^2 \\ \vdots & \vdots & & \vdots \\ p_1^{m-1} & p_2^{m-1} & \cdots & p_n^{m-1} \end{bmatrix}. \tag{4}$$

B. Parameter Estimation

Consider the $m \times 1$ data vector $\{y(t)\}$ as given by Equation (1). In general we first filter the data set $\{y(t)\}$ before decimating, to minimize effects of aliasing (we discuss filtering in detail in Section IV). Here, we consider an lth order FIR filter of the form

$$y'(t) = Ky(t), \tag{5}$$

where

$$K = \begin{bmatrix} k_l & k_{l-1} & k_{l-2} & \cdots & k_0 & 0 & 0 & \cdots & 0 \\ 0 & k_l & k_{l-1} & \cdots & k_1 & k_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & k_l & k_{l-1} & \cdots & k_1 & k_0 & 0 \\ 0 & \cdots & 0 & 0 & k_l & \cdots & k_2 & k_1 & k_0 \end{bmatrix}_{(m-l)\times m}, \tag{6}$$

and where the sequence $\{k_c\}_{c=0}^{l}$ is the FIR filter impulse response.

From $y'_q(t)$ we now define a set of decimated sequences as

$$y'^u_q(t) = y'_{qd+u}(t) \qquad q = 0, 1, \ldots, m'_d - 1 \qquad u = 0, 1, \ldots, d-1, \tag{7}$$

where $m'_d = \lfloor \frac{m-l}{d} \rfloor$. The index "$u$" gives the start sample in the decimation; thus, the sequences $\left\{y'^u_q(t)\right\}$ (for fixed t) represent a set of interleaved sequences decimated from $\{y'(t)\}$. From Equations (2) and (7), we see that each sequence $\left\{y'^u_q(t)\right\}$ is a noisy exponential sequence of the form

$$y'^u_q(t) = \sum_{i=1}^{n} x'^u_i(t)\,(p'_i)^q + e''^u_q(t), \tag{8}$$

where

$$\begin{aligned} p'_i &= (p_i)^d \\ x'^u_i(t) &= x_i(t) p_i^{(u+l)} \mathcal{K}(p_i), \end{aligned} \tag{9}$$

and where $\mathcal{K}(z)$ is the FIR filter polynomial given by

$$\mathcal{K}(z) = k_0 + k_1 z^{-1} + k_2 z^{-2} + \cdots + k_l z^{-l}. \tag{10}$$

Note that the effects of the FIR filtering in the new model are to scale the amplitude coefficients and to color the noise. In general, we will choose K to be a bandpass filter. By careful choice of the FIR filter, we can significantly reduce the mode amplitude coefficients outside of some band of interest; in this case we can assume the number of the "significant" modes in the filtered data is n' which is less than n. In this case, we have the following model:

$$y_q'^u(t) = \sum_{i=1}^{n'} x_i'^u(t) \left(p_i'\right)^q + e_q'^u(t), \tag{11}$$

where

$$e_q'^u(t) = \sum_{i=n'+1}^{n} x_i'^u(t) \left(p_i'\right)^q + e_q''^u(t). \tag{12}$$

Note that $e_q'^u(t)$ is colored Gaussian noise with nonzero mean; the effect of the nonzero mean is to introduce some bias in the parameter estimates, as we will see in the following sections.

Thus we now have $N \times d$ sequences with common poles but different amplitude coefficients. This case is similar to the original multi-snapshot nondecimated model in Equation (2). As a result, the TLS-Prony algorithm can be applied to the data in Equation (7) to give estimates $\left\{\widehat{p_i'}\right\}$ and $\left\{\widehat{x_i'^u}(t)\right\}$. We thus have decimated multi-snapshot backward linear prediction equations given by:

$$\begin{bmatrix} y'^0 & Y'^0 \\ y'^1 & Y'^1 \\ \vdots & \vdots \\ y'^{d-1} & Y'^{d-1} \end{bmatrix} \begin{bmatrix} 1 \\ b' \end{bmatrix} \approx 0, \tag{13}$$

or

$$\begin{bmatrix} y' & Y' \end{bmatrix} \begin{bmatrix} 1 \\ b' \end{bmatrix} \approx 0, \tag{14}$$

where

$$b' = \begin{bmatrix} b'_1 & b'_2 & \cdots & b'_L \end{bmatrix}^T, \tag{15}$$

and each

$$\begin{bmatrix} y'^u & Y'^u \end{bmatrix} = \left[\begin{array}{ccccc} y'_u(1) & y'_{u+d}(1) & y'_{u+2d}(1) & \cdots & y'_{u+Ld}(1) \\ y'_{u+d}(1) & y'_{u+2d}(1) & y'_{u+3d}(1) & \cdots & y'_{u+(L+1)d}(1) \\ \vdots & \vdots & \vdots & & \vdots \\ y'_{u+m'_d-(L+1)d}(1) & y'_{u+m'_d-Ld}(1) & y'_{u+m'_d-(L-1)d}(1) & \cdots & y'_{u+m'_d-d}(1) \\ \hline y'_u(2) & y'_{u+d}(2) & y'_{u+2d}(2) & \cdots & y'_{u+Ld}(2) \\ y'_{u+d}(2) & y'_{u+2d}(2) & y'_{u+3d}(2) & \cdots & y'_{u+(L+1)d}(2) \\ \vdots & \vdots & \vdots & & \vdots \\ y'_{u+m'_d-(L+1)d}(2) & y'_{u+m'_d-Ld}(2) & y'_{u+m'_d-(L-1)d}(2) & \cdots & y'_{u+m'_d-d}(2) \\ \hline & & \vdots & & \\ \hline y'_u(N) & y'_{u+d}(N) & y'_{u+2d}(N) & \cdots & y'_{u+Ld}(N) \\ y'_{u+d}(N) & y'_{u+2d}(N) & y'_{u+3d}(N) & \cdots & y'_{u+(L+1)d}(N) \\ \vdots & \vdots & \vdots & & \vdots \\ y'_{u+m'_d-(L+1)d}(N) & y'_{u+m'_d-Ld}(N) & y'_{u+m'_d-(L-1)d}(N) & \cdots & y'_{u+m'_d-d}(N) \end{array}\right]. \tag{16}$$

Here L is the order of prediction, and b' is the coefficient vector of the polynomial $B'(z)$ given by

$$B'(z) = 1 + b'_1 z + b'_2 z^2 + \cdots + b'_L z^L. \tag{17}$$

The choice of L affects the accuracy of the b'_i coefficients, as we address in later

sections.

The TLS-Prony method considers the effect of noise perturbation of both Y' and y', and the TLS solution attempts to minimize the effect of these perturbations on the prediction coefficient vector b' (see [1] for details). This is accomplished by obtaining an SVD of the matrix $\left[\begin{array}{c:c} y' & Y' \end{array}\right]$ and truncating all but the first n' singular values to arrive at an estimate $\left[\begin{array}{c:c} \widehat{y'} & \widehat{Y'} \end{array}\right]$[1]. Inserting $\left[\begin{array}{c:c} \widehat{y'} & \widehat{Y'} \end{array}\right]$ in Equation (14) gives the modified linear prediction equation

$$\widehat{Y'}\widehat{b'} \approx -\widehat{y'} \tag{18}$$

from which the linear prediction coefficient vector estimate $\widehat{b'}$ is found as

$$\widehat{b'} = -\widehat{Y'}^{+}\widehat{y'}, \tag{19}$$

where $^{+}$ denotes the Moore-Penrose pseudoinverse. Finally, the estimates for the decimated poles are found by

$$\widehat{p'_j} = \mathrm{zero}_j\left(\widehat{B'}(z)\right), \qquad j = 1, 2, \ldots, L. \tag{20}$$

It is not in general possible to recover $\widehat{p_j}$ from $\widehat{p'_j}$, as the mapping $p_i \rightarrow p_i^d$ is not one-to-one. However, the mapping can be made one-to-one by suitably restricting the domain of p_i. For example, if it is known *a priori* that $\angle p_i \in \left(-\frac{\pi}{d}, \frac{\pi}{d}\right)$ then $\widehat{p_j}$ may be uniquely recovered from $\widehat{p'_j}$ by

$$\widehat{p_j} = \left(\widehat{p'_j}\right)^{\frac{1}{d}}. \tag{21}$$

In order to meet the domain restriction requirement without *a priori* information, one can choose a suitable FIR filter K, as discussed in Section IV below.

Once the poles are found, the corresponding amplitude coefficients can be estimated from the decimated pole estimates and decimated data using Equations (11) and (9) or from the nondecimated pole estimates and nondecimated data in precisely the same way as in the nondecimated TLS-Prony solution [10]. Using the decimated pole estimates and decimated data is more computationally efficient, since there will be shorter data lengths and fewer pole estimates. For this case, Equation (11) leads to the following equation for the amplitude coefficients,

$$\widehat{A'}_L \widehat{\mathcal{K}}_p \widehat{X} = Y_a'^0, \tag{22}$$

where

$$\widehat{A'}_L = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \widehat{p_1'} & \widehat{p_2'} & \cdots & \widehat{p_L'} \\ \widehat{p_1'}^2 & \widehat{p_2'}^2 & \cdots & \widehat{p_L'}^2 \\ \vdots & \vdots & & \vdots \\ \widehat{p_1'}^{(m_d'-1)} & \widehat{p_2'}^{(m_d'-1)} & \cdots & \widehat{p_L'}^{(m_d'-1)} \end{bmatrix}$$

$$\widehat{\mathcal{K}}_p = \operatorname{diag}\left(\widehat{p_1}^l \mathcal{K}(\widehat{p_1}), \widehat{p_2}^l \mathcal{K}(\widehat{p_2}), \ldots, \widehat{p_L}^l \mathcal{K}(\widehat{p_L})\right)$$

$$\widehat{X} = \begin{bmatrix} \hat{x}_1(1) & \hat{x}_1(2) & \cdots & \hat{x}_1(N) \\ \hat{x}_2(1) & \hat{x}_2(2) & \cdots & \hat{x}_2(N) \\ \vdots & \vdots & & \vdots \\ \hat{x}_L(1) & \hat{x}_L(2) & \cdots & \hat{x}_L(N) \end{bmatrix}$$

$$Y_a'^0 = \left[\begin{array}{cccc} y'^0(1) & y'^0(2) & \cdots & y'^0(N) \end{array} \right]. \tag{23}$$

The amplitude coefficients can be found from a least squares solution to Equation (22),

$$\widehat{X} = \widehat{K}_p^{-1} \left(\widehat{A'}_L^* \widehat{A'}_L \right)^{-1} \widehat{A'}_L^* Y_a'^0, \tag{24}$$

where * denotes complex conjugate transpose. We note that Equation (24) is not used in practice to solve Equation (23), as more numerically sound procedures (such as QR decomposition [11]) can be used.

Because only n' singular values of $\widehat{Y'}$ are nonzero, there are at most n' pole estimates which can correspond to true data modes. Therefore, only the n' poles which have the largest energy are retained [10]. We then re-estimate the amplitude coefficients of these n' poles by eliminating all but the n' "high energy pole" columns of $\widehat{A'}_L$, then recomputing the least squares solution for $\widehat{X}$. We note that the second amplitude coefficient estimation can be done by using the QR decomposition from the first amplitude coefficient estimation. By doing so we save computation in the second amplitude coefficient estimation. We also note that the noise $e_q'^u(t)$ in Equation (11) is not in general white so an (unweighted) least solution to Equation (22) may not lead to amplitude estimates with smallest variance.

The above procedure can be modified further to provide even greater computational savings. The extra interleaving decimated data sets can be discarded for the computation of the poles, *i.e.*, u in Equation (7) can take on only the value 0. From a Nyquist theory point of view, we note that discarding the interleaving data should provide no loss in performance if an ideal lowpass filter can be implemented (this is

confirmed in the examples presented in Section V below). Note that this data discarding can be easily incorporated into the matrix K by keeping only every dth row and that the statistical analysis below is presented in a general K framework, so that the analysis applies equally well to the cases where data is and is not discarded.

III. Statistical Analysis

A major contribution of this paper is the derivation of the statistical properties of the TLS-Prony pole estimates obtained using decimation. Below we derive a general expression for the first order approximation of the probability distribution function (pdf) of the estimates of $\{p_i'\}$ under the assumptions that there is a filter present as described by Equation (5). This expression applies to different decimation values so it can be used to determine the relative statistical accuracy for various choices in the TLS-Prony algorithm. The expression is given in the following theorem.

Theorem 4.1: Assume we are given FIR filtered data $\{y_q'(t)\}$ as defined in Equations (7) and (11). Let

$$\widehat{P'} = \left[\begin{array}{cccc} \widehat{p_1'} & \widehat{p_2'} & \cdots & \widehat{p_{n'}'} \end{array} \right]^T \tag{25}$$

be the n' highest energy TLS-Prony pole estimates found from Equations (20) and (24). Then the first order approximation (as $\sigma \rightarrow 0$) of the pdf of $\widehat{P'}$ is given by

$$\widehat{P'} \sim \mathcal{N}\left(P' + P_b', \Sigma'\right), \tag{26}$$

where the estimate bias and covariance are given by

$$\begin{aligned} P'_b &= F'G'S'^{+}P'_\epsilon, \\ \Sigma' &= \sigma F'G'S'^{+}\Sigma_\epsilon \left(F'G'S'^{+}\right)^*. \end{aligned} \tag{27}$$

Here P'_ϵ is defined in Equation (68) and Σ_ϵ is given by

$$\Sigma_\epsilon = \begin{bmatrix} \Sigma_\epsilon^0 & \Sigma_\epsilon^1 & \cdots & \Sigma_\epsilon^{d-1} \\ \Sigma_\epsilon^{1*} & \Sigma_\epsilon^0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \Sigma_\epsilon^1 \\ \Sigma_\epsilon^{d-1*} & \cdots & \Sigma_\epsilon^{1*} & \Sigma_\epsilon^0 \end{bmatrix}, \tag{28}$$

where each $\Sigma_\epsilon^{(v-u)}$ is a block diagonal matrix given by

$$\Sigma_\epsilon^{(v-u)} = \operatorname{diag}\left(B'K^uK^{v*}B'^*, B'K^uK^{v*}B'^*, \ldots, B'K^uK^{v*}B'^*\right)_{N(m'_d-L)\times N(m'_d-L)}. \tag{29}$$

The expressions for K^u and B' are given by

$$K^u = \begin{bmatrix} K(u+1)^T & K(u+1+d)^T & \cdots & K(u+1+m'-d)^T \end{bmatrix}^T, \tag{30}$$

where $K(a)$ is the ath row of K, and

$$B' = \begin{bmatrix} 1 & b'_1 & b'_2 & \cdots & b'_L & 0 & 0 & \cdots & 0 \\ 0 & 1 & b'_1 & \cdots & b'_{L-1} & b'_L & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & b'_1 & \cdots & b'_{L-1} & b'_L & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & b'_{L-2} & b'_{L-1} & b'_L \end{bmatrix}_{(m'_d-L)\times(m'_d)} \tag{31}$$

We also have

$$F' = \mathrm{diag}\left(\frac{1}{\eta'_1}, \frac{1}{\eta'_2}, \ldots, \frac{1}{\eta'_{n'}}\right), \tag{32}$$

$$\eta'_i = \begin{bmatrix} b'_1 & b'_2 & \cdots & b'_L \end{bmatrix} \begin{bmatrix} 1 \\ 2p'_i \\ \vdots \\ Lp_i'^{(L-1)} \end{bmatrix}, \tag{33}$$

$$G' = \begin{bmatrix} p'_1 & p_1'^2 & \cdots & p_1'^L \\ p'_2 & p_2'^2 & \cdots & p_2'^L \\ \vdots & \vdots & & \vdots \\ p_{n'} & p_{n'}'^2 & \cdots & p_{n'}'^L \end{bmatrix}, \tag{34}$$

and S' defined as the noise free version of Y'.

Proof: See Appendix. □

The above theorem gives the accuracy expressions for the case where the data are decimated by a factor of d, and all d interleaved data sets are used in the estimation (see Equation (13)). In some cases one may wish to use only one of the d interleaved data sets, in which case Equation (13) is replaced by

$$\begin{bmatrix} y'^0 & Y'^0 \end{bmatrix} \begin{bmatrix} 1 \\ b' \end{bmatrix} \approx 0, \tag{35}$$

where y'^0 and Y'^0 are defined in Equation (16). For this case, the accuracy expressions for the resulting parameter estimates are given in the following corollary.

Corollary 4.1: Assume we are given data as in *Theorem 4.1*, but form TLS-Prony estimates based on Equation (35) instead of Equation (13). Then the first order

approximation (as $\sigma \to 0$) of the pdf of $\widehat{P'}$ is given by

$$\widehat{P'} \sim \mathcal{N}\left(P' + P'_{b1}, \Sigma'_1\right), \tag{36}$$

where

$$\begin{aligned} P'_{b1} &= F'G'S'^{0+}P'^{0}_{\epsilon}, \\ \Sigma'_1 &= \sigma F'G'S'^{0+}\Sigma^0_{\epsilon}\left(F'G'S'^{0+}\right)^*. \end{aligned} \tag{37}$$

Here P'^{0}_{ϵ} is the first block of P'_{ϵ} (the block involves K^0), and Σ^0_{ϵ}, F', and G' are the same as the ones in *Theorem 4.1*, and S'^{0} is the noise free version of Y'^{0}.

Proof: Follow by replacing S' by S'^{0} and K by K^0 in the proof of *Theorem 4.1*. □

Equations (27) and (37) provide the biases and covariances for decimated pole estimates given a particular set of poles and decimation factor. If the nondecimated pole estimates $\{\widehat{p_j}\}$ are recovered from the decimated pole estimates $\left\{\widehat{p'_j}\right\}$ using Equation (21), the biases and variances for the nondecimated poles can be derived in terms of the biases and variances for the decimated poles using a first order approximation of Equation (21). Defining $\widetilde{p'_i} = \widehat{p'_i} - p'_i$. We have the following derivation.

$$\begin{aligned} p'_i + \widetilde{p'_i} &= (\widehat{p}_i)^d \\ &= (p_i + \widetilde{p}_i)^d \\ &= p_i^d + dp_i^{d-1}\widetilde{p}_i + (\text{higher order terms}), \end{aligned} \tag{38}$$

and note that $p'_i = p_i^d$. Thus,

$$\widetilde{p}_i \approx \frac{\widetilde{p'_i}}{dp_i^{d-1}}. \tag{39}$$

Therefore, we have

$$\text{Bias}(\hat{p}_i) \triangleq E[\hat{p}_i] - p_i = \frac{\text{Bias}\left(\hat{p}_i'\right)}{dp_i^{d-1}}$$

$$\text{Var}(\hat{p}_i) \triangleq E\left[(\hat{p}_i - E[\hat{p}_i])(\hat{p}_i - E[\hat{p}_i])^*\right] = \frac{\text{Var}\left(\hat{p}_i'\right)}{d^2 |p_i|^{2(d-1)}}. \quad (40)$$

The variances of the nondecimated poles can now be compared to their respective CRBs (provided the estimated bias is negligible). Recently, a CRB formulation has been developed for multi-snapshot damped exponentials in noise [12]. These CRB results can be directly compared with the variances of the estimated poles using the TLS-Prony method to examine its performance in both nondecimated and decimated circumstances. This comparison is shown in Section V for a number of examples.

IV. Full Spectrum Estimation Using Filtering and Decimation

Using the decimation scheme which has been developed, any poles or modes not in the band of interest need to be filtered out so that they are not aliased into the band of interest by the decimation operation. Even if there are no poles outside the band of interest, a filter can be still applied to reduce the out-of-band noise which will be aliased into the band of interest by the decimation operation. However, the imperfections of a FIR filter, nonideal stopband rejection and data length reduction by transient response effects, will cause loss in performance, as shown in Section V below. In general, nonideal stopband rejection increases bias in the estimation (because the leakage of the stopband poles, and thus the first term in Equation (12), will be larger),

and data length reduction due to transient effects of the FIR filter causes a variance increase.

In this section, we develop a procedure to obtain full spectrum mode estimates by use of bandwidth segmentation, and then discuss how to design the needed FIR filters so that the performance loss is minimized. Finally, we present operation counts for both the decimated and nondecimated estimation procedures.

A. Procedure

To examine the use of filtering and decimation, assume we are interested in estimating modes in the complete spectrum (that is, poles which may lie anywhere in the complex plane), and we wish to use a decimation factor of d. We estimate poles in each region of Figure 1(a) using a decimation based TLS-Prony procedure. First, we modulate the data to center the band of interest about $f = 0$ as follows

$$y_q(t) \rightarrow y_q(t)e^{-j2\pi f_0 q}, \tag{41}$$

where f_0 is the modulating frequency for one of the subbands of interest, $f_0 = 0, \frac{1}{d}, \frac{2}{d}, \ldots, \frac{d-1}{d}$. We then lowpass filter the modulated data to isolate the frequency band $f \in \left[-\frac{1}{2d}, \frac{1}{2d}\right]$. Finally, we apply the decimated TLS-Prony algorithm of Section II. The resulting pole estimates, $\widehat{p'_j}$, as given by Equation (21) lie in $f \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ as shown in Figure 1(b). The corresponding pole estimates in nondecimated frequency space are given by Equation (21) with modulation as follows

$$\widehat{p}_j = \left(\widehat{p'_j}\right)^{\frac{1}{d}} e^{j2\pi f_0}. \tag{42}$$

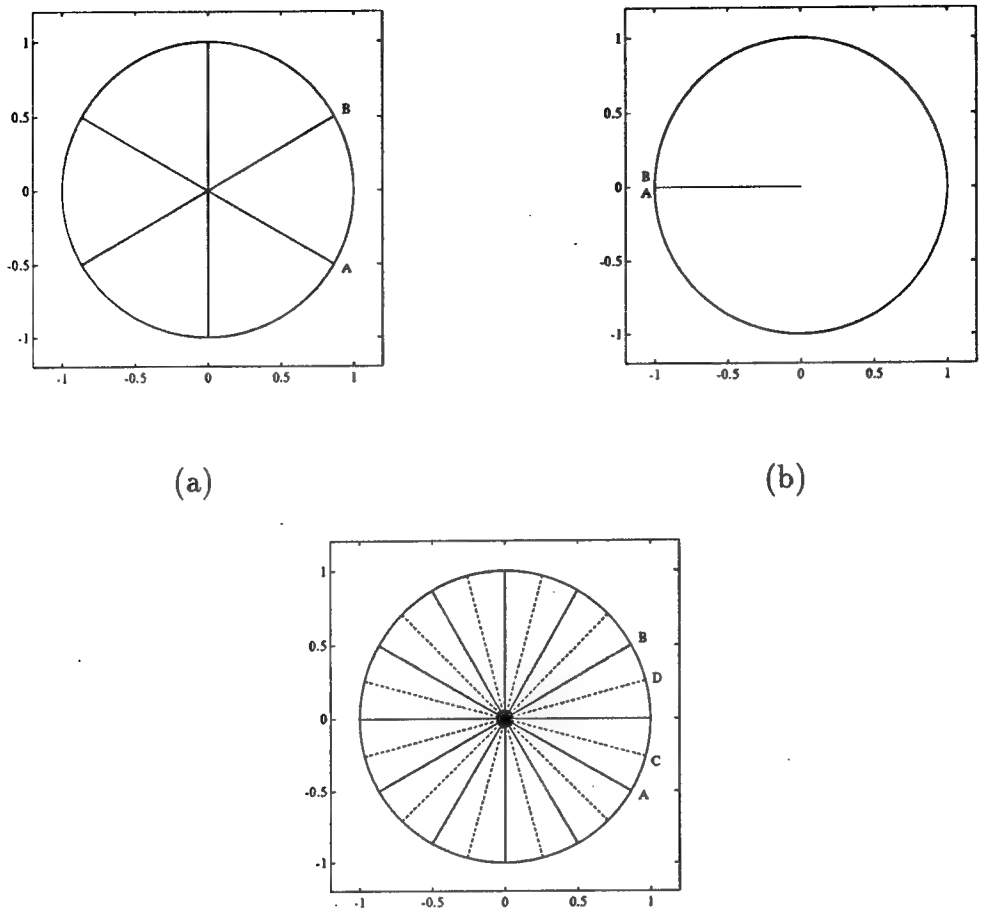


Figure 1: Spectral effect due to decimating by six.

A problem which can result from the above procedure is that poles near the endpoints of the subband region may be incorrectly estimated near the opposite endpoint. This results from the discontinuity of the mapping in Equation (42) for $\angle \widehat{p'_j} \approx \pm\pi$. It can be seen from Figure 1(b) that small errors in estimates near A and B result in large differences in Figure 1(a). To avoid this problem, we use

$c > d$ overlapped estimations, each of size $\frac{1}{d}$ (note that this changes the modulating frequencies to be $f_0 = 0, \frac{1}{c}, \frac{2}{c}, \ldots, \frac{c-1}{c}$). This is shown in Figure 1(c) for $c = 2d$. For each region, we retain only those pole estimates which are in the half of the overlap region which is closer to the center of the band of interest. For example, in Figure 1(c) for the subband [A,B] we retain poles only in the region [C,D]. This corresponds to retaining pole estimates whose angles satisfy

$$\angle \widehat{p_j'} \in \left[-\frac{\pi d}{c}, \frac{\pi d}{c} \right] \tag{43}$$

in the decimated frequency space. In this way, we reduce the effects of the discontinuity of the mapping in Equation (21) for $\angle \widehat{p_j}$ near π. The overlap method also helps to provide immunity to effects of a nonideal lowpass filter, as is discussed in the next subsection.

B. Filter Design and Performance Loss

As we discussed before, the use of finite length FIR filters results in performance loss. We need to design a filter such that this performance loss is minimized. Because the stopband of the filter is set by the decimation factor, d, to be $\left[-\frac{1}{2d}, \frac{1}{2d} \right]$, there remain three free design parameters: the order of the FIR filter (l), the number of overlapped estimations (c), and the flatness of the passband. Note that the passband is defined as $\left[-\frac{1}{2c}, \frac{1}{2c} \right]$, the region in which we actually retain the pole estimates.

Each of the parameters has its own effect in the estimation performance. Larger filter lengths result in a variance increase because of the data length reduction associ-

ated with discarding the transient response of the filter output. Shorter filter lengths result in increased bias of the pole estimates due to aliasing of imperfectly attentuated modes in the stopband region. The number of overlapped estimations primarily influences the computational aspect. A larger c results in more computations; nevertheless, a larger c allows for a larger filter transition band, thus allowing us to design for a better filter stopband. Note that we can tolerate some non-flatness since we can compensate for this effect in the amplitude coefficient estimation. However, a non-flat passband increases the variance of the poles near the minima of the passband since the filter reduces the power of these poles more than other poles in the passband.

After filtering and decimating the data, the stopband signal poles are aliases into the band of interest. If the energy of the filtered stopband signal modes is small relative to the in-band Gaussian noise, the estimation bias caused by the stopband signal poles will be negligible in comparison to the estimation variance. This implies that for high SNR signals we need more stopband rejection in order to avoid the bias problem. One should choose a filter length sufficient to attenuate out-of-band modes to below the noise floor.

A procedure to design the FIR filter is presented here to minimize variance and bias. First, based on SNR one can determine the needed stopband rejection to ensure the leakage of the out-of-band signal poles is small relative to the in-band noise. Then, one can choose the filter order l and the number of the overlapped estimations c to give the desired stopband rejection with a small transition band. Note that we want l and c to be as small as possible. For instance, in the example given below, we

estimated according to the SNR that a 20dB stopband rejection is enough to avoid the bias problem. After few trials, we found that using $c = 2d = 12$ and $l = 20$, an equiripple FIR filter gives the approximated stopband rejection. Compromise was made to sacrifice the flatness of the passband a little since we can compensate its effect during the amplitude coefficient estimation.

C. Operation Counts

The main goal of applying the decimation procedure is a reduction in computations. In this section we compare the computations of a full TLS-Prony estimation applied to a nondecimated sequence to a set of those from a set of TLS-Prony estimates obtained from a set of bandpass filtered and decimated sequences. We include the computations associated with overlapping as well as those associated with the filtering operation for this comparison.

In the operation counts which follow, we assume that only one of the interleaved sets of decimated data is used in the decimation-based TLS-Prony algorithm (as in Equation (35)). In addition, we assume $L = \frac{m}{3d}$ is used for the prediction order, because this prediction order gives near optimal accuracy (see [10, 13] and Section V below). We compute operation accounts for the SVD operation, the QR decomposition, the polynomial root finding operation, and the filtering operation; these were found to account for over 90% of the total operations in our computer simulations.

1. SVD Operation Counts

For a real matrix which has dimension $R \times C$, the approximate floating point operation (flop) count associated with the "economical" SVD computation (in which only the first C left singular vectors are computed) is given by $\mathrm{fc}^{\mathrm{SVD}} \approx 14RC^2 + 8C^3$ [11]. For a complex matrix the count is about a factor of 2 larger. In our case, the matrix $\left[\begin{array}{cc} \hat{y}' & \widehat{Y'} \end{array}\right]$ has dimension $R \times C = \left(\frac{2(m-l)}{3d}N\right) \times \left(\frac{(m-l)}{3d} + 1\right)$. Therefore, for the nondecimated case ($d = 1$ and $l = 0$) we thus get an SVD flop count of

$$\mathrm{fc}^{\mathrm{SVD}}_{\mathrm{nondec}} \approx 28\left(\frac{2}{3}mN\right)\left(\frac{m}{3}+1\right)^2 + 16\left(\frac{m}{3}+1\right)^3, \tag{44}$$

and for the decimated case with c overlap regions we obtain

$$\mathrm{fc}^{\mathrm{SVD}}_{\mathrm{dec}} \approx c\left(28\left(\frac{2(m-l)}{3d}N\right)\left(\frac{m-l}{3d}+1\right)^2 + 16\left(\frac{m-l}{3d}+1\right)^3\right). \tag{45}$$

2. QR Decomposition Operation Counts

For a real matrix, the approximate flop count associated with the QR decomposition is given by $\mathrm{fc}^{\mathrm{QR}} \approx 2RC^2 - \frac{2}{3}C^3$ [11]. For complex matrices, the count is to be a factor of 4 larger. In our case, $R \times C = \left(\frac{m-l}{d}\right) \times \left(\frac{m-l}{3d}\right)$. We thus obtain for the nondecimated and decimated cases, respectively

$$\mathrm{fc}^{\mathrm{QR}}_{\mathrm{nondec}} \approx \frac{8m^3}{9} - \frac{8m^3}{81}, \tag{46}$$

and

$$\mathrm{fc}^{\mathrm{QR}}_{\mathrm{dec}} \approx c\left(\frac{8(m-l)^3}{9d^3} - \frac{8(m-l)^3}{81d^3}\right). \tag{47}$$

3. Polynomial Root Finding Operation Counts

For a real polynomial of order κ, the approximate flop count associated with the root finding operation is given by $\mathrm{fc}^{\mathrm{root}} \approx \frac{20}{3}\kappa^3$ [11]. For complex data considered the count is to be a factor of about 10 larger; in our case $\kappa = \left(\frac{m-l}{3d}\right)$. So we have

$$\mathrm{fc}^{\mathrm{root}}_{\mathrm{nondec}} \approx \frac{200}{81}m^3, \tag{48}$$

and

$$\mathrm{fc}^{\mathrm{root}}_{\mathrm{dec}} \approx c\left(\frac{200(m-l)^3}{81d^3}\right). \tag{49}$$

4. Filtering Operation Counts

For the decimated case, we must also include the operation count associated with the filtering operation. From Equations (5) and (6), the approximate flop count associated with the matrix multiplication induced by this filter is thus given by

$$\mathrm{fc}^{\mathrm{filter}}_{\mathrm{dec}} \approx c\left(\frac{m-l}{d}(l+1)\right) \times 6. \tag{50}$$

The factor of 6 arises because there are 6 flops (4 multiplies and 2 adds) per complex multiplication.

V. Examples and Simulation Studies

Examples using the statistical analysis results are presented here which demonstrate the advantages of using decimation. Simulations are also presented for full

spectrum data sets to demonstrate the estimation ability of the modified TLS-Prony method developed in Section IV. Again, note that we are considering the case where the extra interleaved data sets are discarded.

A. Single Undamped Mode

In this example we assume one snapshot of data of length $m = 140$. We assume a single exponential located on the unit circle and an SNR of 5dB. We compare estimates of this exponential using no decimation and using decimation by $d = 6$. Using the filter design procedure outlined above, we obtain an FIR filter whose frequency response is shown in Figure 2. Here we used $c = 2d$ overlapped estimations and an equiripple FIR filter of order $l = 20$. Note that we could decrease the order of the FIR filter to achieve the same stopband rejection, but in doing so we will obtain a less flat passband, which results in increased variance. Note also that in this case we do not have out-of-band signal poles, so there will be no bias in the pole estimates (to a first order).

Figure 3 shows the theoretical variance of the estimated pole versus prediction order for various decimation factors as compared to the CRB. From this figure we can see that the minimum variance occurs at a lower prediction order for higher d. The minimum occurs for a prediction order equal to about one third of the decimated data length (*i.e.*, $L = \frac{m'}{3}$), which is consistent with results for nondecimated data [10, 13]. This shows that the best performance for the decimated cases occurs at lower prediction orders than for the nondecimated case, thus reducing the computational

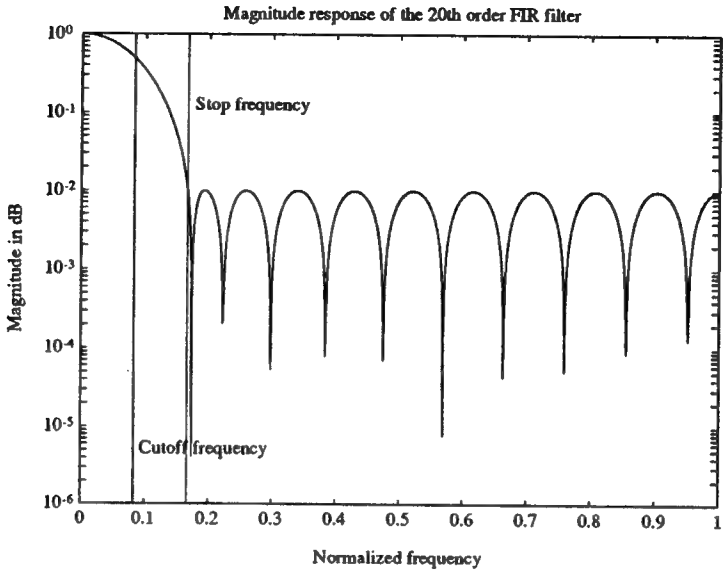


Figure 2: Frequency response of the $20th$ order equiripple FIR filter.

load.

We note, however, that since the data has to be filtered prior to decimation, the curves for the decimated cases peak at about 1.5dB lower than for the nondecimated case for various decimation factors. The performance loss is due to the fact that the transient response portion of the filter output (20 points for this case) needs to be discarded. Note that for a fixed length FIR filter this performance degradation becomes smaller as the data length is increased, since the percent difference between the original and filtered data lengths decreases.

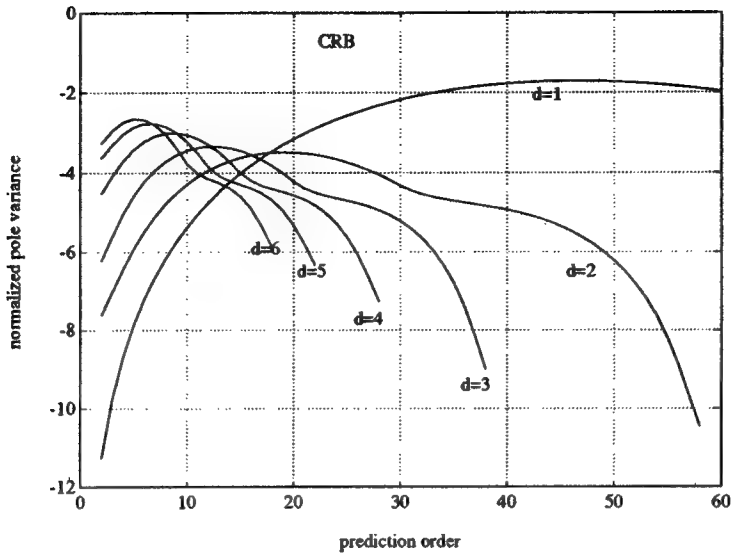


Figure 3: $10\log_{10}(\mathrm{CRB}/\mathrm{Var}(p))$ for various d and prediction orders for a single undamped mode.

B. Two Undamped Modes

In a second example we make the same assumptions above, except that there are two equal energy exponentials located on the unit circle one Fourier bin apart (*i.e.*, $\Delta f = \frac{1}{m} = \frac{1}{140}$). The total SNR is assumed to be 8dB in this case in order to maintain 5dB SNR/pole. Figure 4 shows the theoretical variance of the estimates for one of the poles (the variance of the other pole is similar). We can see that the characteristics are much the same as in the one pole case, the difference being higher variances due to the presence of each pole's neighbor.

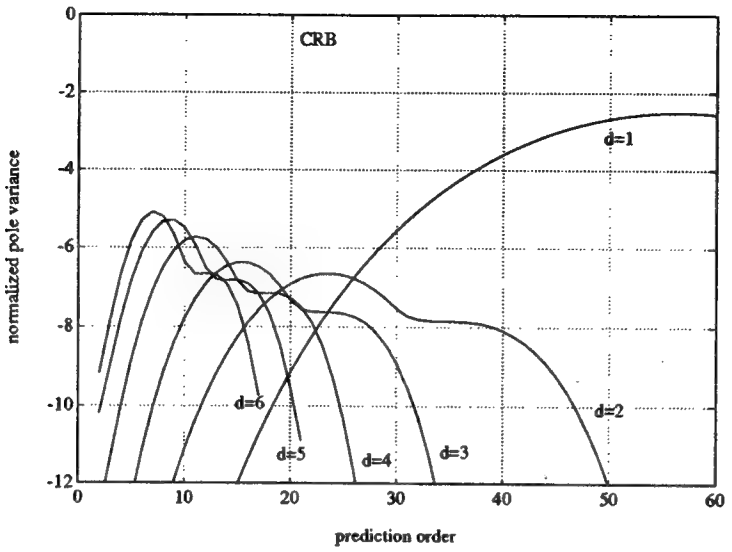


Figure 4: $10\log_{10}(\mathrm{CRB}/\mathrm{Var}(p_1))$ for various d and prediction orders for two undamped modes.

C. *Monte-Carlo Simulation of an Undamped Ten Mode Case*

We now present a set of simulations for a general ten undamped mode case. In these simulations we have $N = 1$ snapshot, and $n = 10$ poles present in the data. The amplitude coefficients all have unit magnitude; the phases of the amplitude coefficients are chosen randomly. We consider two data lengths, $m = 140$ and $m = 560$ data points. Figure 5 shows the locations of the ten poles; each is indicated by an "x". Five-hundred independent Monte-Carlo simulations are performed by adding noise to the data such that the total SNR is 20dB (10dB/pole). Estimates for the poles are obtained using the TLS-Prony algorithm without decimation (*i.e.*, $d = 1$) and with decimation using a decimation factor of $d = 6$.

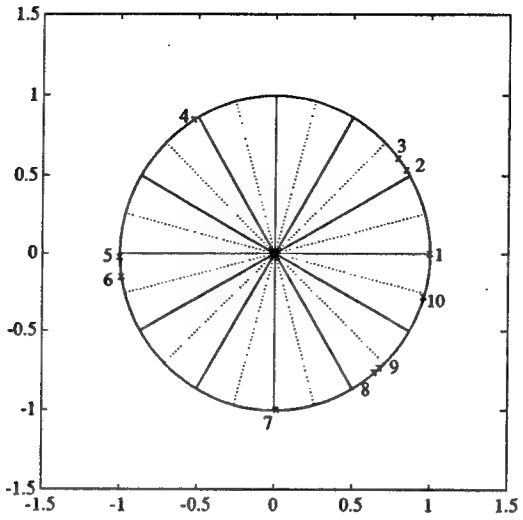


Figure 5: True pole locations for an undamped ten mode case.

For the decimation results, the FIR filter is the same as the one used in the previous examples and the frequency response was shown in Figure 2 (thus, $c = 12$ and $l = 20$). The prediction order is $L = \lfloor \frac{m-l}{3d} \rfloor$; the numbers of singular values retained in the simulations are 10 for $d = 1$ and the number of poles in each subestimation section for $d = 6$ ($\{2, 3, 2, 0, 1, 1, 2, 2, 0, 1, 3, 3\}$ for this case). The prediction orders used correspond to one third of the effective data lengths in the two cases as was suggested by Examples 1 and 2. Prior to the calculation of the amplitude coefficients, poles with magnitude larger than 1.15 and smaller than $\frac{1}{1.15}$ are eliminated to avoid poor conditioning in the least squares solution of the amplitude coefficients.

Table 1: Theoretical and simulation variances, and MSEs for the undamped ten poles, $m = 140$ data point case (all values are in dB).

pole number	CRB	$d = 1$			$d = 6$			
		Theory	Sim.	Sim. MSE	Theory	Sim.	Theory MSE	Sim. MSE
1	-63.4	-61.8	-61.5	-61.5	-60.5	-60.6	-60.3	-60.4
2	-63.0	-60.7	-60.9	-60.8	-59.7	-58.9	-59.5	-58.7
3	-63.0	-60.7	-61.0	-60.9	-60.4	-58.7	-59.9	-58.0
4	-63.6	-61.4	-61.2	-61.2	-60.9	-61.0	-60.9	-60.9
5	-63.4	-62.3	-62.3	-62.2	-60.6	-60.4	-60.1	-59.9
6	-63.4	-62.3	-62.2	-62.2	-60.7	-60.5	-60.4	-60.3
7	-63.5	-61.4	-61.1	-61.1	-61.0	-60.8	-60.9	-60.7
8	-55.1	-51.4	-51.0	-50.9	-47.3	-48.4	-47.0	-48.1
9	-55.1	-51.4	-51.3	-51.1	-46.8	-47.9	-46.5	-47.4
10	-63.4	-61.8	-61.6	-61.6	-59.6	-59.6	-59.5	-59.6

1. Performance Comparison

Tables 1–3 summarize the performance of the various methods for this example.

We first consider Table 1 which shows the $m = 140$ data point case. From Table 1 we see that the Monte-Carlo variances for the $d = 1$ case were 1.2dB to 4.1dB away from their CRBs. Note that the Monte-Carlo variances are within 0.5dB of those predicted by the theory which substantiates the theory (the differences are due to

the fact that theory is only a first order analysis). Note also that the estimates have negligible bias which is shown by the fact that the pole variances are very close to their MSEs.

The Monte-Carlo variances for the $d = 6$ case were 2.6dB to 7.2dB away from their CRBs. This represents an average 1.7dB loss for the entire $d = 6$ system versus the nondecimated system. The main cause of the performance loss is the 20 data point loss due to the transient response of the FIR filter outputs. Note that the Monte-Carlo variances are within 1.7dB of statistical theory. By comparing the simulation variances and MSEs, we see that some of pole estimates are slightly biased. When decimation is used, *Theorem 4.1* gives an analytical expression for the this bias; theoretical biases are compared to biases obtained from Monte-Carlo simulations with good agreement in most cases.

If the data length is increased to $m = 560$ points, the theoretical and simulation results show even better agreement, as is shown in Table 3. In this case, the overall loss using decimation is less than 0.4dB compared to the nondecimated case. In addition, the simulation variances are within 0.4dB of the theoretically derived variances, and the bias of the pole estimates is significantly reduced.

2. Operation Count Comparison

The $d = 1$ and $d = 6$ estimation procedures are now compared on the basis of their computational costs. We only compare the results for $m = 140$ data points. Using MATLAB, the "economical" version of the SVD operation, the left division opera-

Table 2: Theoretical and simulation biases for the poles ($m = 140$, unit= 10^{-3})

pole number	Theory	Sim.
1	0.078-0.179i	0.088-0.196i
2	0.166-0.033i	0.255-0.037i
3	0.330+0.072i	0.449+0.188i
4	0.042+0.085i	0.005+0.100i
5	-0.082-0.314i	-0.074-0.324i
6	0.086+0.205i	0.018+0.184i
7	0.076-0.077i	0.067-0.136i
8	1.251-0.171i	0.847-0.693i
9	-1.121-0.470	-1.134-0.625i
10	-0.095+0.142i	-0.047+0.021i

tion (using QR decomposition to solve least squares problems), and the root finding operations required an average of 16.5Mflops for each of the Monte-Carlo simulations for the $d = 1$ case. Each of the twelve $d = 6$ SVDs, QR decompositions, polynomial root findings, and filtering operations required an average of 53.7Kflops, resulting in a total of 644.4Kflops for each Monte-Carlo simulation. The computational savings for the SVDs, QR decompositions, polynomial root findings, and filtering operations in this example using decimation was thus a factor of about 25.6. This compares with a savings factor of 24.0 which is predicted by Equations (44)–(50) for this scenario.

Table 3: Theoretical and simulation variances, and MSEs for the poles for the 560 data point case (all values are in dB).

pole		$d = 1$			$d = 6$			
number	CRB	Theory	Sim.	Sim. MSE	Theory	Sim.	Theory MSE	Sim. MSE
1	-81.6	-79.9	-79.6	-79.6	-79.7	-79.8	-79.7	-79.8
2	-81.6	-79.6	-79.7	-79.7	-79.4	-79.4	-79.4	-79.3
3	-81.6	-79.6	-79.7	-79.7	-79.3	-79.4	-79.3	-79.4
4	-81.7	-79.9	-79.7	-79.6	-79.7	-79.7	-78.9	-79.0
5	-81.6	-79.8	-79.8	-79.8	-79.5	-79.2	-79.4	-79.1
6	-81.6	-79.8	-79.7	-79.7	-79.4	-79.4	-79.1	-78.9
7	-81.7	-79.9	-79.9	-79.8	-79.7	-79.4	-79.6	-79.3
8	-81.6	-79.9	-80.0	-80.0	-79.6	-79.4	-79.0	-79.0
9	-81.6	-80.0	-80.0	-80.0	-79.5	-79.4	-79.5	-79.4
10	-81.7	-80.0	-80.1	-80.1	-79.3	-79.7	-79.3	-79.7

Note that these numbers cannot be computed directly in part due to differences between MATLAB's SVD, QR decomposition, and root finding algorithms and the ones used for operation counts in Equations (44)–(50). The average total flop counts (including all operations) for the nondecimated and decimated Monte-Carlo simulations were 17.4Mflops and 698Kflops, respectively, to give a savings factor of 24.9. Note that the four computational cost components we have detailed make up about 92% of the total computations. With higher decimation factors the savings are even more

substantial.

VI. Conclusions

In this paper we have developed a TLS-Prony estimation algorithm which incorporates data decimation. We also have developed a statistical analysis for estimated poles of this algorithm. We have shown through examples using this analysis that decimation provides a minimum variance for estimated poles that occurs at a prediction order which is smaller than the optimal prediction order for nondecimated data by a factor of d, thus allowing for computational savings. We have shown that this benefit is obtained at the expense of pole variance performance due to the filtering which is required; this expense becomes smaller for longer data lengths. We have also shown how the modified TLS-Prony method can be used on full spectrum data one band at a time to realize the computational savings in a more general signal framework.

With this decimation procedure, we are now able to make a well-quantified tradeoff of accuracy for computation when using the TLS-Prony estimation procedure.

Appendix: Proof of Theorem

From Equation (14) we can make the following substitutions

$$\left(S' + \tilde{S}'\right)\left(b' + \tilde{b}'\right) = -\left(s' + \tilde{s}'\right), \tag{51}$$

where s' is the noise free version of y', $\widehat{Y'} = S' + \tilde{S}'$, $\widehat{y'} = s' + \tilde{s}'$, and $\widehat{b'} = b' + \tilde{b}'$. We can see that the $\tilde{}$ terms are small perturbations for the high SNR case, which is assumed. Multiplying out Equation (51) and retaining only the first order $\tilde{}$ terms gives[1]

$$S'b' + \tilde{S}'b' + S'\tilde{b}' = -s' - \tilde{s}'. \tag{52}$$

Now note that $S'b' = -s'$ since they are the noiseless terms. Equation (52) thus becomes

$$S'\tilde{b}' = -\left(\tilde{s}' + \tilde{S}'b'\right). \tag{53}$$

Multiplying both sides through by $S'S'^{+}$ and noting that $S'S'^{+}S' = S'$, we obtain

$$S'\tilde{b}' = -S'S'^{+}\left(\tilde{s}' + \tilde{S}'b'\right). \tag{54}$$

Let $Y' = S' + W'$ and $y' = s' + w'$, where W' and w' are the appropriate noise matrices (*i.e.*, of the form given in Equation (14) and composed of the noise sequences in Equation (12)). Thus we can see that $\tilde{S}'$ and $\tilde{s}'$ are perturbations caused by W', w', and the SVD truncation. By using perturbation analysis [14] on the matrices $\left[\begin{array}{cc} s' & S' \end{array}\right]$, $\left[\begin{array}{cc} y' & Y' \end{array}\right]$, and $\left[\begin{array}{cc} \widehat{y'} & \widehat{Y'} \end{array}\right]$, it can be shown that to first order approximation

$$S'^{+}\left[\begin{array}{cc} \tilde{s}' & \tilde{S}' \end{array}\right] = S'^{+}\left[\begin{array}{cc} w' & W' \end{array}\right]. \tag{55}$$

Thus, Equation (54) can be written as

$$S'\tilde{b}' = -S'S'^{+}\epsilon, \tag{56}$$

[1]Note that the approximation is valid since the matrices $\left[\begin{array}{c:c} \widehat{y'} & \widehat{Y'} \end{array}\right]$ and $\left[\begin{array}{c:c} s' & S' \end{array}\right]$ have the same rank.

where $\epsilon = w' + W'b'$.

Observing the data model and the formulation of the S' matrix, we can write S'^u as

$$S'^u = \left[\begin{array}{cccc}
x_1(1)p_1^u & x_2(1)p_2^u & \cdots & x_{n'}(1)p_n^u \\
x_1(1)p_1^{u+d} & x_2(1)p_2^{u+d} & \cdots & x_{n'}(1)p_{n'}^{u+d} \\
\vdots & \vdots & & \vdots \\
x_1(1)p_1^{u+m-(L+1)d} & x_2(1)p_2^{u+m-(L+1)d} & \cdots & x_{n'}(1)p_{n'}^{u+m-(L+1)d} \\
\hline
x_1(2)p_1^u & x_2(2)p_2^u & \cdots & x_{n'}(2)p_n^u \\
x_1(2)p_1^{u+d} & x_2(2)p_2^{u+d} & \cdots & x_{n'}(2)p_{n'}^{u+d} \\
\vdots & \vdots & & \vdots \\
x_1(2)p_1^{u+m-(L+1)d} & x_2(2)p_2^{u+m-(L+1)d} & \cdots & x_{n'}(2)p_{n'}^{u+m-(L+1)d} \\
\hline
 & \vdots & & \\
\hline
x_1(N)p_1^u & x_2(N)p_2^u & \cdots & x_{n'}(N)p_n^u \\
x_1(N)p_1^{u+d} & x_2(N)p_2^{u+d} & \cdots & x_{n'}(N)p_{n'}^{u+d} \\
\vdots & \vdots & & \vdots \\
x_1(N)p_1^{u+m-(L+1)d} & x_2(N)p_2^{u+m-(L+1)d} & \cdots & x_{n'}(N)p_{n'}^{u+m-(L+1)d}
\end{array}\right] G', \qquad (57)$$

or

$$S'^u = H^u G'. \qquad (58)$$

Letting $H = \left[\begin{array}{cccc} H^{0T} & H^{1T} & \cdots & H^{d-1T} \end{array}\right]^T$, we get

$$S' = HG'. \qquad (59)$$

Equation (56) thus becomes

$$HG'\tilde{b}' = -HG'S'^{+}\epsilon. \qquad (60)$$

Now note by definition that the true and estimated Lth order characteristic polynomials are $B(z) = 1 + b'_1 z + b'_2 z^2 + \cdots + b'_L z^L$ and $\widehat{B'}(z) = 1 + \hat{b}'_1 z + \hat{b}'_2 z^2 + \cdots + \hat{b}'_L z^L$, respectively. Hence $B\left(p'_i\right) = 0$ and $\widehat{B'}\left(\hat{p}'_i\right) = 0$.

We can use a first order Taylor expansion to find an expression for the error in the estimated pole locations. For each $\hat{p}'_i$ we obtain

$$
\begin{aligned}
0 &= \widehat{B'}\left(\hat{p}'_i\right) \\
&= \widehat{B'}\left(p'_i\right) + \frac{\partial}{\partial z}\widehat{B'}(z)|_{z=p'_i}\left(\hat{p}'_i - p'_i\right) + (\text{higher order terms}) \\
&= \widehat{B'}\left(p'_i\right) - B'\left(p'_i\right) + \frac{\partial}{\partial z}\widehat{B'}(z)|_{z=p'_i}\left(\hat{p}'_i - p'_i\right) + ((\text{higher order terms}) \\
&\approx 1 + \hat{b}'_1 p'_i + \hat{b}'_2 p_i'^2 + \cdots + \widehat{b'_L} p_i'^L - \left(1 + b'_1 p'_i + b'_2 p_i'^2 + \cdots + b'_L p_i'^L\right) \\
&\quad + \left(\hat{b}'_1 + 2\hat{b}'_2 p'_i + \cdots + L\widehat{b'_L} p_i'^{(L-1)}\right)\left(\hat{p}'_i - p'_i\right) \\
&\approx \left[\begin{array}{cccc} p'_i & p_i'^2 & \cdots & p_i'^L \end{array}\right]\left[\begin{array}{c} \hat{b}'_1 - b'_1 \\ \hat{b}'_2 - b'_2 \\ \vdots \\ \widehat{b'_L} - b'_L \end{array}\right] + \left[\begin{array}{cccc} b'_1 & b'_2 & \cdots & b'_L \end{array}\right]\left[\begin{array}{c} 1 \\ 2p'_i \\ \vdots \\ Lp_i'^{(L-1)} \end{array}\right]\left(\hat{p}'_i - p'_i\right) \\
&= \left[\begin{array}{cccc} p'_i & p_i'^2 & \cdots & p_i'^L \end{array}\right]\left(\hat{b}' - b'\right) + \eta'_i\left(\hat{p}'_i - p'_i\right), \qquad (61)
\end{aligned}
$$

or, to first order,

$$\left(\hat{p}'_i - p'_i\right) = -\frac{1}{\eta'_i}\left[\begin{array}{cccc} p'_i & p_i'^2 & \cdots & p_i'^L \end{array}\right]\tilde{b}'. \qquad (62)$$

Thus, for all of the n' true poles we obtain

$$\widehat{P'} - P' = -F'G'\tilde{b}'. \qquad (63)$$

Since H is full rank (this can be seen by noting that each block of rows is simply

a Vandermonde matrix derived from distinct poles times a diagonal matrix of the nonzero amplitude coefficients), we can multiply Equation (60) by $(H^*H)^{-1}H^*$ to get

$$G'\tilde{b}' = -G'S'^{+}\epsilon, \tag{64}$$

and, by substituting Equation (64) into Equation (63), we obtain

$$\widehat{P'} - P' = F'G'S'^{+}\epsilon. \tag{65}$$

We now note that $\begin{bmatrix} w' & W' \end{bmatrix}\begin{bmatrix} 1 \\ b' \end{bmatrix}$ can be written as

$$\begin{bmatrix} w' & W' \end{bmatrix}\begin{bmatrix} 1 \\ b' \end{bmatrix} = \left[\begin{array}{c} B'e'^{0}(1) \\ B'e'^{0}(2) \\ \vdots \\ B'e'^{0}(N) \\ \hline B'e'^{1}(1) \\ B'e'^{1}(2) \\ \vdots \\ B'e'^{1}(N) \\ \hline \vdots \\ \hline B'e'^{d-1}(1) \\ B'e'^{d-1}(2) \\ \vdots \\ B'e'^{d-1}(N) \end{array}\right] = \left[\begin{array}{c} B'K^{0}(s^{r}(1)+e(1)) \\ B'K^{0}(s^{r}(2)+e(2)) \\ \vdots \\ B'K^{0}(s^{r}(N)+e(N)) \\ \hline B'K^{1}(s^{r}(1)+e(1)) \\ B'K^{1}(s^{r}(2)+e(2)) \\ \vdots \\ B'K^{1}(s^{r}(N)+e(N)) \\ \hline \vdots \\ \hline B'K^{d-1}(s^{r}(1)+e(1)) \\ B'K^{d-1}(s^{r}(2)+e(2)) \\ \vdots \\ B'K^{d-1}(s^{r}(N)+e(N)) \end{array}\right], \tag{66}$$

where

$$s^r(t) = \left[\sum_{i=n'+1}^{n} x_i(t)\,(p_i)^0 \quad \sum_{i=n'+1}^{n} x_i(t)\,(p_i)^1 \quad \cdots \quad \sum_{i=n'+1}^{n} x_i(t)\,(p_i)^{m-1} \right]^T, \quad (67)$$

and where B' and K^u are given by Equations (30) and (31). Recall $\{e(t)\}$ are zero mean Gaussian and thus ϵ is multivariate Gaussian with mean

$$E[\epsilon] = P'_\epsilon, \quad (68)$$

where P'_ϵ is in the form of Equation (66) without the $e(t)$'s, and covariance matrix

$$\begin{aligned} \mathrm{Cov}(\epsilon) &= E\left[\left(\begin{bmatrix} w' & W' \end{bmatrix}\begin{bmatrix} 1 \\ b' \end{bmatrix}\right)\left(\begin{bmatrix} w' & W' \end{bmatrix}\begin{bmatrix} 1 \\ b' \end{bmatrix}\right)^*\right] \\ &= \sigma\Sigma_\epsilon, \end{aligned} \quad (69)$$

where Σ_ϵ is given by Equation (28).

Equations (65), (68) and (69) imply that the mean and covariance matrix of $\widehat{P'}$ are given by Equation (27). □

References

[1] M. A. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–35, no. 10, pp. 1440–1454, October 1987.

[2] W. M. Steedly and R. L. Moses, "High resolution exponential modeling of fully polarized radar returns," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-27, no. 3, pp. 459–469, May 1991.

[3] B. Liu and F. Mintzer, "Calcuation of narrow-band spectra by direct decimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-26, no. 6, pp. 529–534, December 1978.

[4] M. P. Quirk and B. Liu, "Improving resolution for autoregressive spectral estimation by decimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-31, no. 3, pp. 630–637, June 1983.

[5] M. J. Villalba and B. K. Walker, "Spectrum manipulation for improved resolution," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 6, pp. 820–831, June 1989.

[6] C. J. Ying, R. L. Moses, and R. L. Dilsavor, "Perturbation Analysis for Pole Estimates of Damped Exponential Signals," technical report, The Ohio State University, Department of Electrical Engineering, ElectroScience Laboratory, August 1991.

[7] G. Xu, S. D. Silverstein, R. H. Roy, and T. Kailath, "Parallel implementation and performance analysis of beamspace ESPRIT," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, (Toronto, Ontario), pp. 1497–1500, May 14–17, 1991.

[8] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Development, performance analysis, and experimental evaluation of beamspace root-MUSIC," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, (Toronto, Ontario), pp. 3049–3052, May 14–17, 1991.

[9] S. D. Silverstein, W. E. Engeler, and J. A. Taridif, "Parallel architectures for multirate superresolution spectrum analyzers," *IEEE Transactions on Circuits and Systems*, vol. 38, no. 4, pp. 449–453, April 1991.

[10] W. M. Steedly, C. J. Ying, and R. L. Moses, "Statistical analysis of TLS-based Prony techniques," *Automatica, Special Issue on Statistical Signal Processing and Control.* Accepted.

[11] G. H. Golub and C. F. V. Loan, *Matrix Computations – second edition.* Baltimore: The Johns Hopkins University Press, 1989.

[12] W. M. Steedly and R. L. Moses, "The Cramér-Rao bound for pole and amplitude estimates of damped exponential signals in noise," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, (Toronto, Ontario), pp. 3569–3572, May 14–17, 1991.

[13] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, no. 5, pp. 814–824, May 1990.

[14] Y. Hua and T. K. Sarkar, "A perturbation property of the TLS-LP method," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, no. 11, pp. 2004–2005, November 1990.

C. Reprint of "Resolution Bound and Detection Results for Scattering Centers"

The following pages contain a preprint of the paper, "Resolution Bound and Detection Results for Scattering Centers," in Proceeding of the International Radar 92 Conference, October 12–13, 1992, Brighton, UK, pp. 518–521.

RESOLUTION BOUND AND DETECTION RESULTS FOR SCATTERING CENTERS*

William M. Steedly,[1] Ching-Hui J. Ying,[2] and Randolph L. Moses[2]

[1]The Analytic Sciences Corporation, USA
[2]Department of Electrical Engineering, The Ohio State University, USA

INTRODUCTION

This paper is concerned with detection and estimation of the scattering centers of a target from coherent, stepped frequency measurements. In particular, we are interested in the following questions: 1) how closely spaced can scattering centers be before it is impossible to resolve them, and 2) what is the relationship between the detection probability of a scattering center and the false alarm probability as a function of scattering center SNR.

To address these questions, we hypothesize a parametric model of target scattering. This model assumes the frequency-domain scattering to be a sum of exponential terms. If the exponential terms are undamped, then the model specializes to a point-scatterer assumption. If the exponential terms are not undamped, the model incorporates frequency-dependent radar cross section of scattering centers (see (1)). We consider a particular class of algorithms for estimating the parameters in this exponential model, the so-called total-least squares (TLS) Prony algorithm (2). The TLS-Prony technique has gained popularity as a parameter estimation algorithm for the exponential model because it provides accurate parameter estimates at moderate computational cost (3); it has also been successfully applied to both one-dimensional and two-dimensional radar scattering data (4,5).

Under the exponential model assumption of scattering, the resolution and detection bounds can be reformulated in terms of parameter estimation accuracy for the exponential model. We present the asymptotic (as SNR $\to \infty$) probability density function (pdf) for the exponential model parameter estimates using the TLS-Prony algorithm. We then use this asymptotic pdf to derive scattering center resolution and detection bounds for the TLS-Prony algorithm, and compare these results to Cramér-Rao bound (CRB) results. Monte-Carlo simulations are also presented to compare with the theory. The resolution bounds are obtained from the standard deviation bounds of the pole angles in the exponential models. The detection bounds are obtained by considering the probability that the energy of an estimated mode exceeds a pre-defined threshold. In each case, the probabilities are obtained by considering a high-SNR approximation of the statistical probabilities of exponential model parameter estimates.

One of the advantages of a model-based scattering center estimation procedure is the capability of resolving scattering centers more accurately than is possible using fast Fourier Transform (FFT) techniques. We show that for sufficiently high SNR, both the CRB resolution and the TLS-Prony resolution is better than can be obtained using the FFT.

*THIS RESEARCH WAS SUPPORTED IN PART BY THE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH, THE AVIONICS DIVISION, WRIGHT LABORATORIES, AND THE SURVEILLANCE DIVISION, ROME LABORATORIES.

DATA MODEL AND TLS-PRONY ESTIMATION PROCEDURE

Data Model

Assume the data vector y of length m is modeled as a noisy exponential sequence

$$y_q = \sum_{i=1}^{n} z_i p_i^q + e_q \qquad q = 0,1,\ldots,m-1. \tag{1}$$

There are n distinct exponential modes in the data. Here, it is assumed that $\{e_q\}$ is a zero mean complex white Gaussian noise sequence with variance σ. Equation 1 may be compactly written as

$$y = Az + e, \tag{2}$$

where e $(m \times 1)$ and z $(n \times 1)$ are vectors, and

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ p_1 & p_2 & \cdots & p_n \\ \vdots & \vdots & & \vdots \\ p_1^{m-1} & p_2^{m-1} & \cdots & p_n^{m-1} \end{bmatrix}. \tag{3}$$

TLS-Prony Estimation Procedure

In this subsection we give an overview of the TLS-Prony technique (2) which is used to estimate the parameters of the exponential model presented in Equation 2.

First, Lth order backward linear prediction equations are formed:

$$\begin{bmatrix} y & Y \end{bmatrix} \begin{bmatrix} 1 \\ b \end{bmatrix} \approx 0, \tag{4}$$

where

$$b = \begin{bmatrix} b_1 & b_2 & \cdots & b_L \end{bmatrix}^T \tag{5}$$

and where

$$\begin{bmatrix} y & Y \end{bmatrix} = \begin{bmatrix} y_0 & y_1 & \cdots & y_L \\ y_1 & y_2 & & y_{L+1} \\ \vdots & \vdots & & \vdots \\ y_{m-(L+1)} & y_{m-L} & \cdots & y_{m-1} \end{bmatrix}. \tag{6}$$

In general, $L \geq n$; however, choosing $L > n$ results in more accurate parameter estimates (6).

The solution of Equation 4 involves obtaining an SVD of the matrix $\begin{bmatrix} y & Y \end{bmatrix}$ and truncating all but the first n singular values to arrive at an estimate $\begin{bmatrix} \hat{y} & \hat{Y} \end{bmatrix}$. The linear prediction coefficient vector estimate $\hat{b}$ is found as $\hat{b} = -\hat{Y}^{+}\hat{y}$, where $^{+}$ denotes the Moore-Penrose pseudoinverse. Finally, the pole estimates are found by

$$\hat{p}_j = \text{zero}_j\left(\hat{B}(z)\right), \qquad j = 1,2,\ldots,L. \tag{7}$$

where $\hat{B}(z) = 1 + \hat{b}_1 z + \cdots + \hat{b}_L z^L$. Once the poles have been determined, the amplitude coefficients can be found

as the least squares solution to

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \widehat{p}_1 & \widehat{p}_2 & \cdots & \widehat{p}_L \\ \vdots & \vdots & & \vdots \\ \widehat{p}_1^{m-1} & \widehat{p}_2^{m-1} & \cdots & \widehat{p}_L^{m-1} \end{bmatrix} \begin{bmatrix} \widehat{z}_1 \\ \widehat{z}_2 \\ \vdots \\ \widehat{z}_L \end{bmatrix} \approx \mathbf{y}. \quad (8)$$

Note that L mode estimates are obtained, of which n are "true" modes. For $L > n$, $L - n$ of the mode estimates are extraneous. If n is known then the true modes can be identified as the n highest energy estimates (3). However, in practice n is typically unknown; in this case the number of singular values retained is a fixed upper bound of n, and true modes could be separated from extraneous modes by using a mode energy threshold (as discussed below).

STATISTICAL ANALYSIS

In order to establish a resolution bound and probabilities of detection and false alarm, we need the statistics for the estimated parameters in the TLS-Prony model. For the resolution bounds and detection probabilities, we need only the statistics for the "true" modes (that is, the n modes with highest energies). For this case, the statistics of the parameter estimates have been found in (3,7). However, for the false alarm probability, we also need the statistics of the $L-n$ extraneous modes. The derivation of the extraneous mode statistics appears in (8); the main result is summarised in the following Theorem.

Theorem 1: Assume $\mathbf{y}$ is as given in Equation 1. Let $p = \{p_i\}_{i=1}^{n}$ and $z = \{z_i\}_{i=1}^{n}$ be as in Equation 1, and let $p^e = \{p_i^e\}_{i=1}^{L-n}$ and $z^e = \{z_i^e\}_{i=1}^{L-n}$ denote the $L-n$ extraneous modes obtained in the TLS-Prony procedure when $\sigma = 0$. Let θ denote the $4L \times 1$ vector containing the angles and magnitudes of p, z, p^e, and z^e, respectively, and let $\widehat{\theta}$ denote the TLS-Prony estimate of θ. Then the asymptotic (as $\sigma \to 0$) probability density function of $\widehat{\theta}$ is Gaussian:

$$\widehat{\theta} \sim N(\theta, \sigma \cdot \Sigma_\theta), \quad (9)$$

where Σ_θ is a covariance matrix which depends on m, L, and $\{z_i, p_i\}_{i=1}^{L}$.

The proof of the Theorem and an explicit expression for Σ_θ, can be found in (8). Also in (8) is an expression for Σ_θ when θ is reparameterized in terms of the real and imaginary parts of the poles and amplitude coefficients. We note that the above Theorem gives a theoretical expression for the complete pdf of the estimated parameters; this pdf can then be used to study the resolution properties and detection probabilities of scattering centers, as is discussed below.

A RESOLUTION CRITERION AND BOUND ANALYSIS

Consider two poles on the unit circle, $p_1 = \alpha_1 e^{j\omega_1}$ and $p_2 = \alpha_2 e^{j\omega_2}$. We define the resolution limit r as

$$r = 2\sigma_{\omega_1} + 2\sigma_{\omega_2}, \quad (10)$$

where the angle variances of p_1 and p_2 are $\sigma_{\omega_1}^2$ and $\sigma_{\omega_2}^2$, respectively. When two poles are at this limit, the 95% confidence intervals of the angle estimates for each pole become disjoint.

The CRB resolution limit is found by using CRB expressions for the pole angles in two pole model. The CRBs for the pole angles are inserted into σ_{ω_1} and σ_{ω_2} in Equation 10; such expressions can be found, for example, in (9). This limit gives a lower bound for all unbiased estimators since it is based on the CRB. The resolution limit for the TLS-Prony estimation algorithm is similarly given by using pole angle variance statistics for the true poles; these are readily obtained from $\sigma \cdot \Sigma_\theta$ (see (8)).

Bound Analysis for Two Undamped Modes

In this study we consider two equal energy modes located on the unit circle at $p_1 = e^{j\pi f/m}$ and $p_2 = e^{-j\pi f/m}$ for a data length of $m = 10$, where f is the separation of the two modes in Fourier bins. For the TLS-Prony simulations a prediction order of $L = 4$ was used.

Figure 1 shows the bounds for the CRB, TLS-Prony statistical theory, and TLS-Prony Monte-Carlo results. The axes are normalised to make the curves independent of data length for the CRB curve and the TLS-Prony statistical theory curve. From these curves we can see that the TLS-Prony theoretical bound is quite close to the CRB over a wide range of SNR. Recall that the CRB and TLS-Prony bounds are both derived using small perturbation analysis, so hold only for high SNR. The TLS-Prony Monte Carlo simulations show good agreement with the theoretical bound above 15 dB SNR. We can see that above 20dB SNR/pole/bin the TLS-Prony method virtually achieves the CRB.

Below 20dB SNR/pole/bin, the Monte-Carlo simulations give much higher variance than both the theoretical TLS-Prony curve and the CRB curve. Note, however, that the TLS-Prony analytical variance expression was derived under the assumption of high SNR, and is not expected to be accurate at low SNR. In addition, the CRB is not necessarily a tight bound at low SNR. Thus, it is not clear what the minimum achievable variance is in this region. For example, it is not known whether (or how much) an iterative maximum likelihood procedure would result in lower variance in this region.

We note that above about 18 dB SNR/pole/bin, the TLS-Prony technique gives resolutions of less than one Fourier bin. Therefore, the resolution of the TLS-Prony technique is better than FFT-based techniques since FFT-based techniques can only resolve to within one Fourier bin. If windowing is used in conjunction with the FFT-based methods, their resolution is even larger than one Fourier bin (*e.g.*, it would be about 1.8 Fourier bins using a Hamming window).

DERIVATION OF PROBABILITIES OF DETECTION AND FALSE ALARM

In practice, one does not know *a priori* how many scattering centers are present. In this case, one would accept or reject a mode estimate as a scattering center based on some threshold. We consider a threshold on the energy of the estimated exponential mode, as this corresponds to radar cross section of an estimated scattering center.

In the TLS-Prony method, one obtains estimates of L poles and L corresponding amplitude coefficients. From this, one can compute the energy E_j of each of the L modes by

$$E_j = \beta_j^2 \sum_{i=0}^{m-1} \alpha_j^{2i}, \quad j = 1, 2, \ldots, L, \quad (11)$$

where β_j and α_j are the magnitudes of the jth amplitude coefficient and pole, respectively. We consider an estimated mode to be detected as a valid scattering center if its energy exceeds a prespecified threshold, and we

reject the mode as an invalid scattering center if it does not. We then can present detection results in the form of receiver operation characteristic (ROC) curves.

Probability of Detection

We define a detection to be the case in which all of the true mode energy estimates exceed an energy threshold, E^0. We thus now derive the energy statistics for the true modes (i.e., $j = 1, 2, \ldots, n$). These statistics can be found from the statistical pdf given in Theorem 1. It can be shown (8) that the energies are noncentral χ^2 distributed; for high SNR, the noncentral χ^2 distribution is well-approximated by a Gaussian distribution. Using first-order approximations, it is possible to derive the mean and covariance of this distribution. Thus, from Theorem 1 we have the following corollary.

Corollary 1: Let

$$E = \left[\begin{array}{cccc} E_1 & E_2 & \cdots & E_n \end{array} \right]^T, \qquad (12)$$

denote the parameter vector for the mode energies of the true modes (i.e., the mode energies corresponding to p and z). Let $\widehat{E}$ denote the estimated energies corresponding to the TLS-Prony parameter estimates. Then the asymptotic (as $\sigma \to 0$) pdf of E is given by

$$\widehat{E} \sim N\left(E, \sigma \cdot \Sigma_E\right), \qquad (13)$$

where Σ_E depends on m, L, and $\{z_i, p_i\}_{i=1}^n$. An explicit expression for Σ_E can be found in (8).

Given the true mode energy distribution, the probabilty of detection, P_D, is given by

$$P_D = \Pr\left(\widehat{E_1} > E^0, \widehat{E_2} > E^0, \ldots, \widehat{E_n} > E^0\right). \qquad (14)$$

This probability is readily computed using Equation 13.

To verify that the theoretical energy distribution given above, Monte-Carlo simulations were performed for a two mode case. In this case, the data consists of two equal energy modes, with $z_1 = z_2 = 1$, located on the unit circle spaced one Fourier bin apart at $p_1 = e^{j2\pi/m}$ and $p_2 = e^{-j2\pi/m}$ for a data length of $m = 10$. A prediction order of $L = 4$ was used, and two singular values were retained. The SNR for these simulations was 10dB/pole.

Figure 2 shows a comparison between the theoretical pdf and a histogram obtained from Monte-Carlo simulations (note that both modes have the same theoretical pdfs and had similar histograms). It can be seen that the theoretical energy distribution is a good approximation to simulation results in this case.

Probability of False Alarm

We define a false alarm to be the case in which one or more of the extraneous modes is above the energy threshold, E^0, and thus misidentified as a true mode. We can derive the statistal properties of the estimated extraneous mode energies in a similar manner as above. In this case, however, the "true" energies of the extraneous modes are zero, so the extraneous modes are distributed as central Chi-squared with two degrees of freedom, χ_2^2 (see (8)). This is stated in the following corollary.

Corollary 2: Let $\left\{\widehat{E_i'}\right\}_{i=1}^{L-n}$ denote the estimated energies corresponding to the TLS-Prony parameter estimates of the extraneous mode energies. Then the asymptotic (as $\sigma \to 0$) pdfs of these energies are given by

$$\widehat{E_i'} \sim \chi_2^2\left(\sigma \cdot \Sigma_{E_i'}\right) \qquad i = 1, 2, \ldots, L-n. \qquad (15)$$

where $\Sigma_{E_i'}$ does not depend on σ. An explicit expression for $\Sigma_{E_i'}$ can be found in (8).

Given the extraneous mode energy distributions, the probabilty of false alarm, P_{FA}, is given by

$$P_{FA} = 1 - \Pr\left(\widehat{E_1'} < E^0, \widehat{E_2'} < E^0, \ldots, \widehat{E_{L-n}'} < E^0\right). \qquad (16)$$

Note that since the extraneous mode energy distribution is Chi-squared with two degrees of freedom, P_{FA} can be evaluated using a Rayleigh distribution.

Figure 3 shows a comparison between the theoretical pdf for the extraneous modes in the previous two modes example and a histogram obtained from Monte-Carlo simulations. Note that the theoretical predictions agree closely with Monte-Carlo simulations.

ROC Analysis for Two Undamped Modes

Using the above detection and false alarm probability results, we can derive ROC curves for scattering center detection at various SNRs. Figure 4 presents such curves for the case considered in Figures 2 and 3, but with varying SNR (note that $1 - P_D$ is actually plotted along the vertical axis). For an SNR at or above 10dB per pole, P_D is always above 0.9 even when P_{FA} is very small (*e.g.*, 10^{-7}). However this is not the case for lower SNR. Note that for low SNRs P_D never reaches one even if P_{FA} is one. This is because the true mode energy distributions were approximated as Gaussian, and the tail of this Gaussian distribution gives a nonzero probability of a negative energy (the noncentral χ_2^2 distribution does not have such a tail). For high SNR, the approximation becomes more valid.

In computing the curves in Figure 4, it was assumed that the extraneous mode energy distributions are independent. Note that this is a worst case assumption, since P_{FA} would decrease if the extraneous mode energies were dependent.

CONCLUSIONS

In the paper we presented resolution bounds and detection results for estimating the scattering centers. These bounds are based on an exponential model of target scattering, which generalizes the point scattering model. The popular TLS-Prony algorithm was used to estimate the parameters of the exponential model. A high SNR statistical analysis of the TLS-Prony algorithm was first presented. Then, based on the results of the statistical analysis, both resolution bounds and detection results were presented. The resolution bounds were compared with both the Cramér-Rao Bound and with Monte-Carlo simulations. It was shown that for an SNR above 18 dB, the TLS-Prony method is capable of resolution to less than a Fourier bin.

The probabilities of the detection and the the false alarm were derived based on the mode energy distributions for both the true and the extraneous modes. For high SNR true mode energy distributions can be approximated as Gaussian distributions, but extraneous mode energy distributions are central chi-squared distributed. These detection and false alarm probabilities can be presented as ROC curves for scattering center detection for examples of interest.

REFERENCES

1. M. P. Hurst and R. Mittra, "Scattering center analysis via Prony's method," *IEEE Trans. Antennas and Propagation*, vol. AP-35, pp. 986–988, Aug. 1987.
2. M. A. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, pp. 1440–1454, Oct. 1987.
3. W. M. Steedly, C. J. Ying, and R. L. Moses, "Statistical analysis of TLS-based Prony techniques," *Automatica*, Submitted for publication.
4. W. M. Steedly and R. L. Moses, "High resolution exponential modeling of fully polarized radar returns," *IEEE Trans. on Aerospace and Electronic Systems*. (to appear).
5. J. J. Sacchini, W. M. Steedly, and R. L. Moses, "Two-dimensional prony modeling and parameter estimation," *IEEE Trans. on Signal Processing*, Submitted for publication.
6. P. Stoica, T. Söderström, and F. Ti, "Asymptotic properties of the high-order Yule-Walker estimates of sinusoidal frequencies," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 1721–1734, Nov. 1989.
7. W. M. Steedly, C. J. Ying, and R. L. Moses, "Statistical analysis of SVD-Based prony techniques," in *Proceedings of the Twenty-Fifth Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA), pp. 232–236, November 4–6 1991.
8. C. J. Ying, W. M. Steedly, and R. L. Moses, "Statistical analysis of true and extraneous tls-prony mode estimates," tech. rep., The Ohio State University, Department of Electrical Engineering, ElectroScience Laboratory, 1992.
9. W. M. Steedly and R. L. Moses, "The Cramér-Rao bound for pole and amplitude estimates of damped exponential signals in noise," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, (Toronto, Ontario), pp. 3569–3572, May 14–17, 1991.

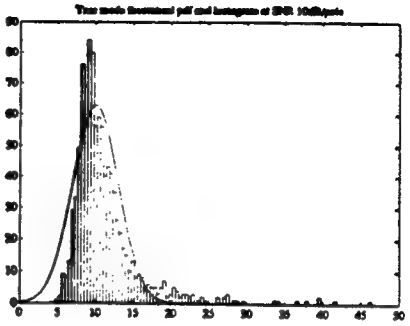


Figure 2: True mode theoretical pdf and histogram for two equal energy undamped modes.

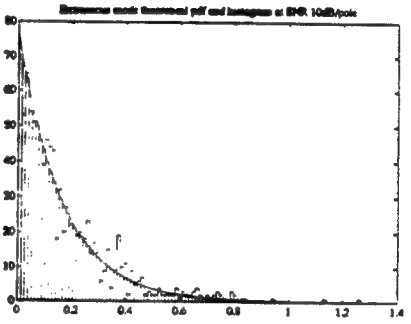


Figure 3: Extraneous mode theoretical pdf and histogram for two equal energy undamped modes.

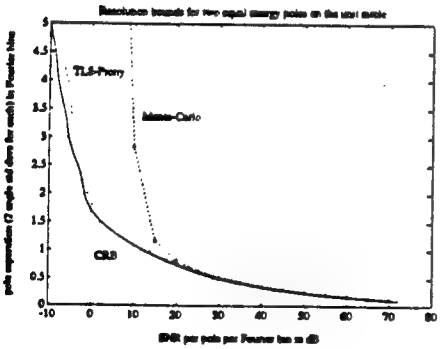


Figure 1: Resolution bounds for two equal energy undamped modes.

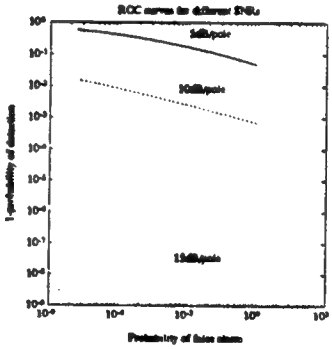


Figure 4: ROC curves for two equal energy undamped modes.

D. Reprint of "Prony Modeling of Linear FM Radar Data"

The following pages contain a preprint of the technical report, "Prony Modeling of Linear FM Radar Data".

Prony Modeling of Linear FM Radar Data

Ching-Hui J. Ying and Randolph L. Moses

Department of Electrical Engineering
The Ohio State University
Columbus, Ohio 43210

Abstract

In this report some preliminary research results on the parametric modeling of linear FM radar data are presented. Superimposed exponential signals are used to model the data. The model parameters are then estimated by the TLS-Prony estimator. In addition to the TLS-Prony algorithm, a modified algorithm incorporating the TLS-Prony estimator and a residue concept is proposed to estimate the parameters. The estimates from the modified algorithm are then improved by a maximum likelihood estimator since the residue algorithm introduces bias on the estimates. The estimation results are compared to the results from the fast Fourier transform technique. The issue of reduced bandwidth is also considered in this report.

I. Introduction

This report presents some preliminary research on the parametric modeling of linear FM radar data. Based on the scattering center assumption the data are modeled by superimposed exponential signals [1],

$$y_q = \sum_{i=1}^{n} x_i p_i^q + e_q \qquad q = 0, 1, \ldots, m-1, \tag{1}$$

where n is the number of the scattering centers and m is the number of frequency measurements. In addition, each p_i is a pole corresponding to a scattering center, each a_i is the amplitude associated with that pole, and e_q is measurement noise.

The problem is to estimate the parameters in the exponential model. Many algorithms have been developed to achieve this objective. One of the most popular methods is the TLS-Prony algorithm [2]. Numerical studies and statistical analysis have shown that for high SNR, the TLS-Prony estimator is unbiased and has very promising performance. Details about the TLS-Prony analysis can be found in [3]. In this report we are concerned about how the TLS-Prony algorithm performs with respect to the traditional fast Fourier transform (FFT) technique, when applied to radar data. For this study we use linear FM radar data of an aircraft that was collected at Rome Laboratories.

In addition, a modified algorithm incorporating TLS-Prony and a residue concept is proposed. The parameter estimates from the modified algorithm then are improved by a maximum likelihood estimator (MLE) since the residue algorithm introduces bias on the parameter estimates. The MLE presented here is a complex version of the MLE presented in [4].

Bandwidth reduction is also considered in the report. Theoretically, parametric modeling has no resolution limit (practically, of course, this is not the case). Here we would like to know how parametric modeling performance degrades as bandwidth is reduced. The TLS-Prony estimation performance with reduced bandwidth is compared to full bandwidth performance of the FFT technique.

An outline of this report is as follows. Section II presents the TLS-Prony algorithm and the modified algorithm. In Section III, the estimated results from both algorithms are compared to each other and to the FFT technique as well. Section IV concludes this report.

II. Estimation Algorithms

A. TLS-Prony Algorithm

Based on the data sequence in Equation (1), backward linear prediction equations are formulated as:

$$\begin{bmatrix} y & : & Y \end{bmatrix} \begin{bmatrix} 1 \\ b \end{bmatrix} \approx 0, \tag{2}$$

where

$$b = \begin{bmatrix} b_1 & b_2 & \cdots & b_L \end{bmatrix}^T, \tag{3}$$

and where

$$\left[\begin{array}{c:c} y & Y \end{array}\right] = \begin{bmatrix} y_0 & y_1 & y_2 & \cdots & y_L \\ y_1 & y_2 & y_3 & \cdots & y_{L+1} \\ \vdots & \vdots & \vdots & & \vdots \\ y_{m-(L+1)} & y_{m-L} & y_{m-(L-1)} & \cdots & y_{m-1} \end{bmatrix}. \tag{4}$$

Here L is the order of prediction and b is the coefficient vector of the polynomial $B(z)$ given by

$$B(z) = 1 + b_1 z + b_2 z^2 + \cdots + b_L z^L. \tag{5}$$

For the noiseless case, L can be any integer greater than or equal to the model order n; however, choosing $L > n$ results in more accurate parameter estimates.

The solution of Equation (2) involves obtaining an SVD of the matrix $\left[\begin{array}{c:c} y & Y \end{array}\right]$ and truncating all but the first n singular values to arrive at an estimate $\left[\begin{array}{c:c} \hat{y} & \hat{Y} \end{array}\right]$ [2]. This leads to the modified linear prediction equation

$$\hat{Y}\hat{b} = -\hat{y} \tag{6}$$

from which the linear prediction coefficient vector estimate $\hat{b}$ is found as

$$\hat{b} = -\hat{Y}^{+}\hat{y}, \tag{7}$$

where $^{+}$ denotes the Moore-Penrose pseudoinverse. Finally, the estimates for the poles are found by

$$\hat{p}_j = \text{zero}_j\left(\hat{B}(z)\right), \qquad j = 1, 2, \ldots, L. \tag{8}$$

Note that for high frequency radar target responses, the scattering is specular and thus the poles are very near the unit circle. That is, the magnitudes of the estimated

poles are close to unity. Before estimating the amplitude coefficients we thus force the estimated poles to be on the unit circle (by setting the magnitude to be one). It is believed this constraint on the poles may give better results. Here we will focus on the constraint case.

Once the poles have been determined, the amplitude coefficients can be found using the pole estimates and Equation (1). This leads to the following least squares equation for the amplitude coefficients,

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \widehat{p}_1 & \widehat{p}_2 & \cdots & \widehat{p}_L \\ \widehat{p}_1^2 & \widehat{p}_2^2 & \cdots & \widehat{p}_L^2 \\ \vdots & \vdots & & \vdots \\ \widehat{p}_1^{m-1} & \widehat{p}_2^{m-1} & \cdots & \widehat{p}_L^{m-1} \end{bmatrix} \begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \\ \vdots \\ \widehat{x}_L \end{bmatrix} \approx \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{m-1} \end{bmatrix} \tag{9}$$

or

$$\widehat{A}_L \widehat{X}_L \approx Y_a. \tag{10}$$

The amplitude coefficients can be found from a least squares solution to Equation (10),

$$\widehat{X}_L = \left(\widehat{A}_L^* \widehat{A}_L\right)^{-1} \widehat{A}_L^* Y_a = \widehat{A}_L^\dagger Y_a, \tag{11}$$

where * denotes complex conjugate transpose. Because only n singular values of $\left[\begin{array}{c:c} \widehat{y} & \widehat{Y} \end{array}\right]$ are nonzero, there are at most n pole estimates which can correspond to data modes. Therefore, only the n poles which have the largest energy are retained. This is done by computing the L mode energies as

$$E_i = \sum_{t=1}^{N} |\widehat{x}_i|^2 \sum_{q=0}^{m-1} |\widehat{p}_i|^{2q} \qquad i = 1, 2, \ldots, L \tag{12}$$

and retaining those n poles whose corresponding energies are highest. We then re-estimate the amplitude coefficients of these n poles. This is done using Equation (11), except that $\hat{A}_L$ is replaced by $\hat{A}$, where $\hat{A}$ is the Vandermonde matrix composed only of the n columns of $\hat{A}_L$ corresponding to the n highest energy poles.

B. A Modified TLS-Prony Algorithm Incorporating a Residue Concept

A modified algorithm incorporating the TLS-Prony algorithm with a residue concept is proposed here. After estimating the parameters via the TLS-Prony algorithm, we can use Equation (1) to create the estimated data sequence. Assuming perfect modeling on the true signals the difference between the measurement data sequence and the estimated data sequence would be the noise sequence only. However, since the modeling is not perfect, the residue contains some information about the data. This motivates the following algorithm.

- Use the estimated parameters from TLS-Prony to create the estimated data sequence via Equation (1). Subtract this sequence from the original sequence, giving a residual data sequence.
- Use the TLS-Prony algorithm on the residual data sequence to obtain residue model parameters. Combine the two sets of parameters (the set from the original measurement data and the set from the residual data).
- Use a maximum likelihood estimator (MLE) to improve the combined estimates.

The MLE is used to overcome bias resulting from estimating the two parameter sets separately. The MLE chosen is the complex version of the MLE presented in [4].

III. Results

In this report a Hamming window is always used to reduced the sidelobes at the cost of losing resolution. The results shown here will only focus on one of the seventeen data sequences since these sequences are responses for the same target, and are nearly identical to one another.

A. Some Comparisons using the data sequence wb0001

In Figure 1 the magnitude of the inverse FFT (IFFT) of the 1024 point data sequence wb0001 using a Hamming window is shown. The high energy section of this IFFT is shown in Figure 2. We will use this result as a standard to compare the estimates from both the constrained TLS-Prony algorithm and the modified TLS-Prony coupled with the residue and MLE algorithm.

The estimation result for the constrained TLS-Prony algorithm using the original data is compared to the IFFT of the original data in Figure 3 (the solid line is for the estimated data and the broken line is the original data). Although the estimated data do not fit the original data very well, from Figure 4, which shows the line spectrum of the TLS-Prony estimates, we can see that at some points (as indicated in the figure) we superresolve closely spaced scattering centers.

For the modified algorithm the results are even better and shown in Figures 5 and 6. In Figure 5 it is clear that the estimated data (solid line) fit the original data (broken line) very well. The line spectrum in Figure 6 shows the superresolvability of the modified algorithm. Many scattering centers not seen in the original data are clear now. This extra resolution is obtained at the cost of increased computation.

In the modified algorithm we use TLS-Prony twice, which doubles the computation. In addition, the computational cost of the MLE has to be added. On average the computation for the modified algorithm is four times that of the computation for the constrained TLS-Prony algorithm.

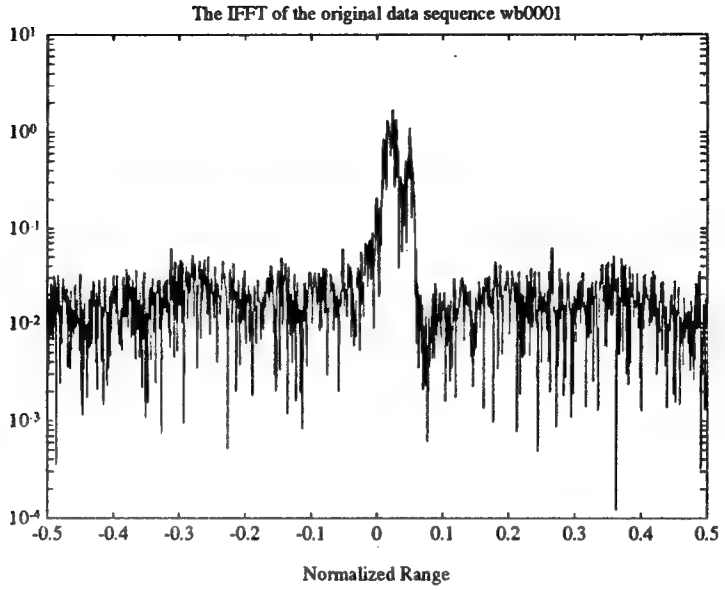


Figure 1: The IFFT of the original data sequence.

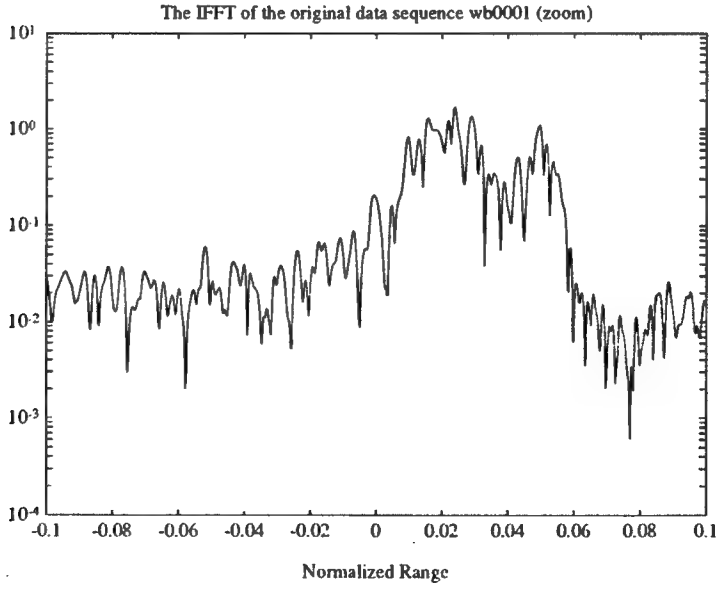


Figure 2: The high energy section of the IFFT of the original data sequence.

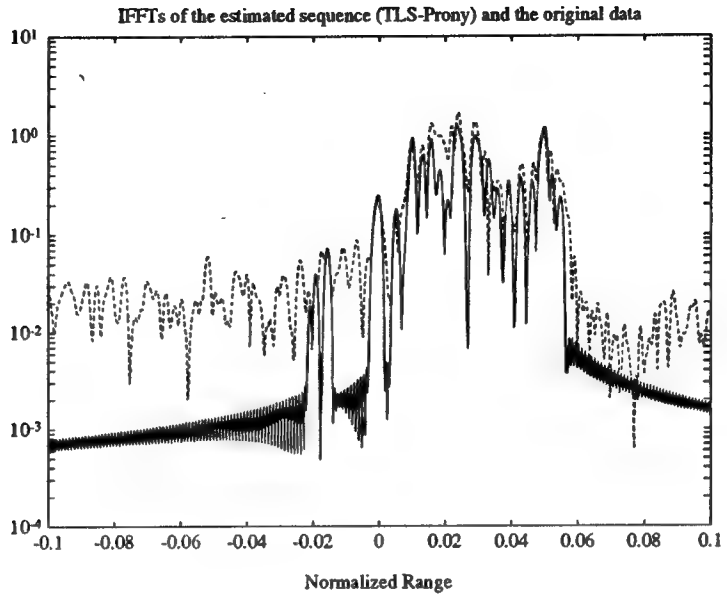


Figure 3: The IFFTs of the original data sequence and the estimated data sequence by the TLS-Prony algorithm using the original data.

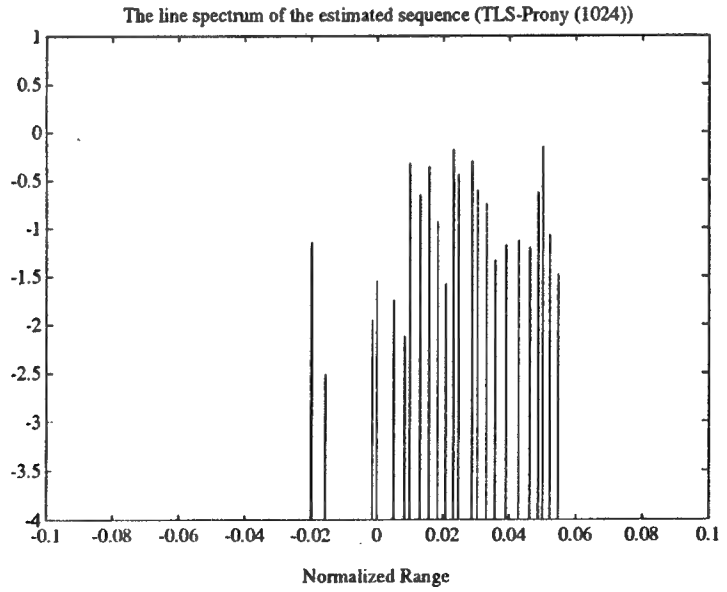


Figure 4: The line spectrum of the TLS-Prony estimates using the original data.

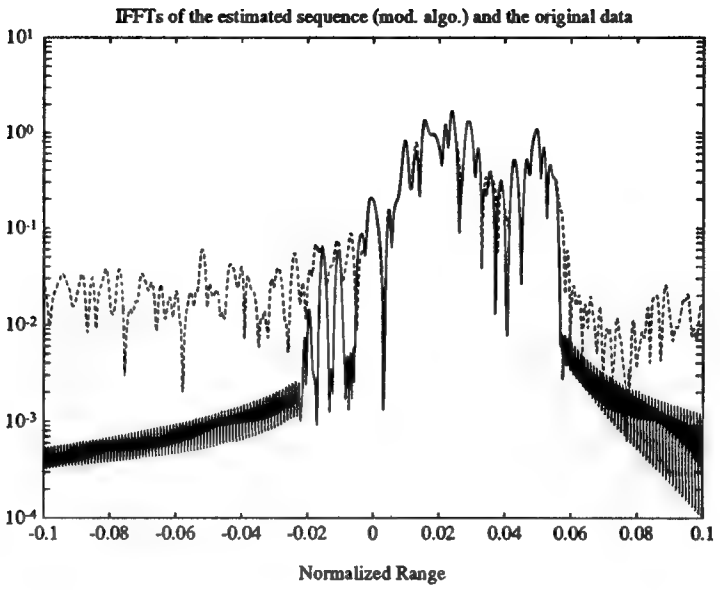


Figure 5: The IFFTs of the original data sequence and the estimated data sequence by the modified algorithm using the original data.

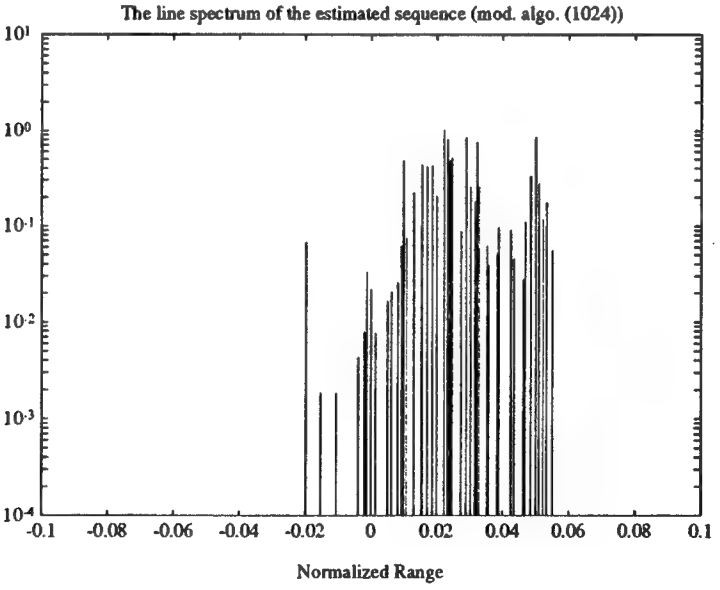


Figure 6: The line spectrum of the modified algorithm estimates using the original data.

B. Reduced bandwidth consideration

In the previous subsection we have shown that the modeling techniques based on the TLS-Prony algorithm give very promising results, and outperform the conventional FFT technique in terms of resolving closely spaced scattering centers. This motivates the following question. If a smaller set of the original data sequence is used, is the performance of these modeling techniques still good enough to compete with the FFT techniques using full bandwidth? We investigate this question using some results based on the same set of data as before.

In Figure 7 the IFFTs of the original data (broken line) and the estimates from the TLS-Prony algorithm using the middle 512 points (solid line) are shown. The corresponding line spectrum is shown below in Figure 8. Here we can see that the estimate is not impressive. In fact, the TLS-Prony algorithm misses some of the scattering centers (as indicated in the figure). However the results from the modified algorithm are much better. They are shown in Figures 9 and 10. These figures are analogous to Figures 7 and 8 except the modified algorithm is used. It is clear from these figures that the modified algorithm obtains all the scattering centers shown for the FFT technique, and it also even superresolves some closely spaced scattering centers. This suggests that the required bandwidth can be reduced by a factor of two when the modified algorithm is used.

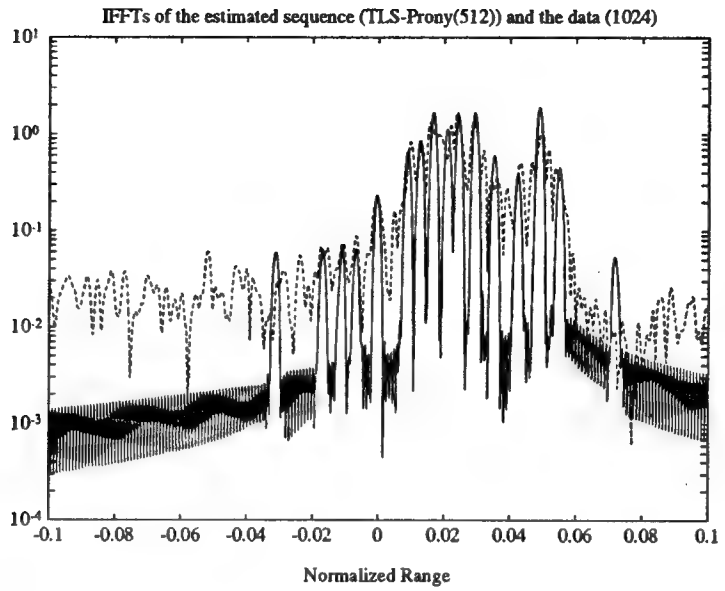


Figure 7: The IFFTs of the original data sequence and the estimated data sequence by the TLS-Prony algorithm using the middle 512 points of the original data .

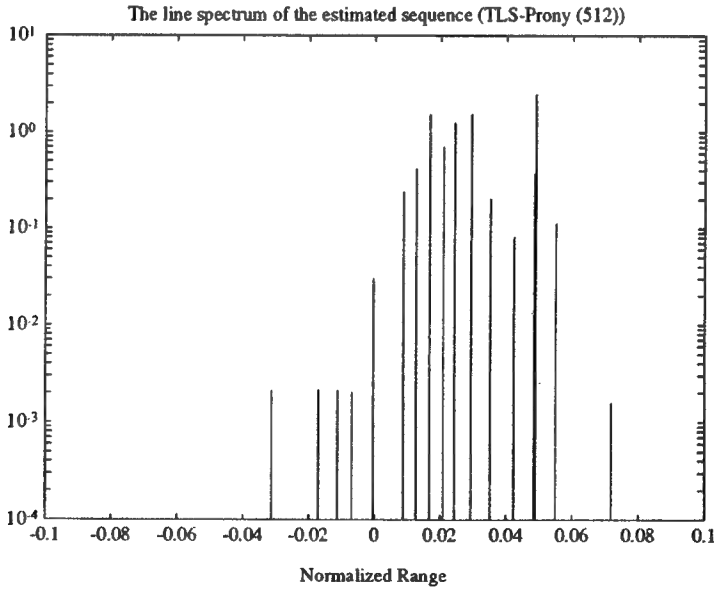


Figure 8: The line spectrum of the TLS-Prony estimates using the middle 512 points of the original data.

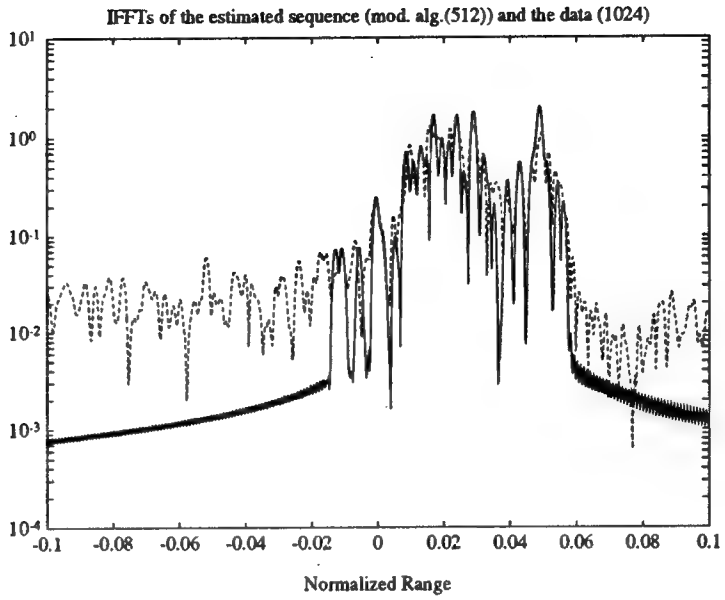


Figure 9: The IFFTs of the original data sequence and the estimated data sequence by the modified algorithm using the middle 512 points of the original data.

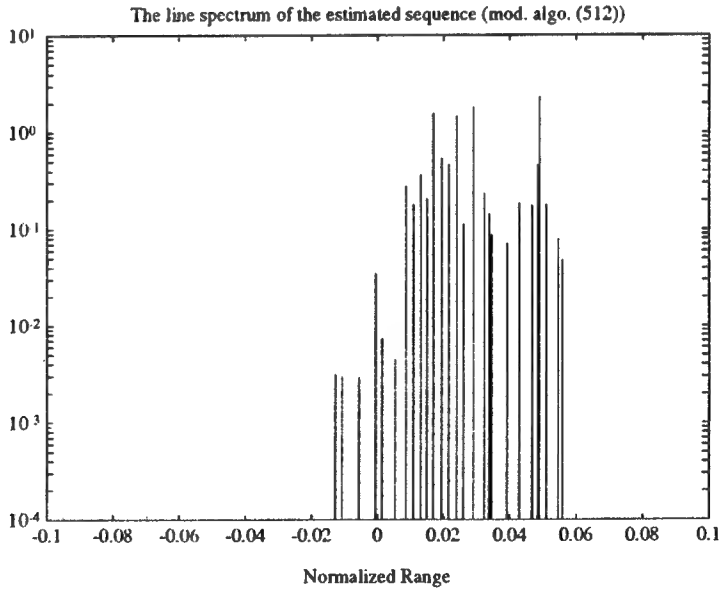


Figure 10: The line spectrum of the modified algorithm estimates using the middle 512 points of the original data.

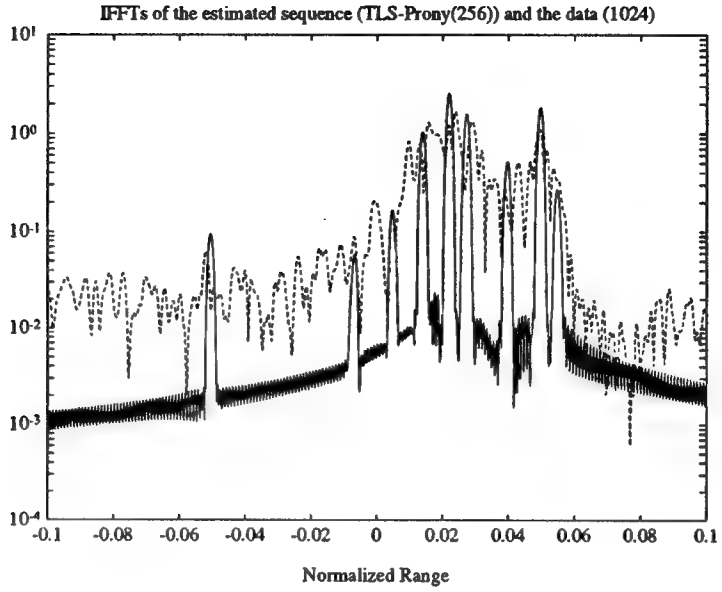


Figure 11: The IFFTs of the original data sequence and the estimated data sequence by the TLS-Prony algorithm using the middle 256 points of the original data.

In Figures 11, 12, 13, and 14 we show the results for using the middle 256 points of the original data sequence. For the TLS-Prony algorithm (in Figures 11 and 12) several scattering centers are missed. For the modified algorithm, however, the results are more encouraging. These results are shown in Figures 13 and 14. Some scattering centers are still missed, but not the stronger scattering centers.

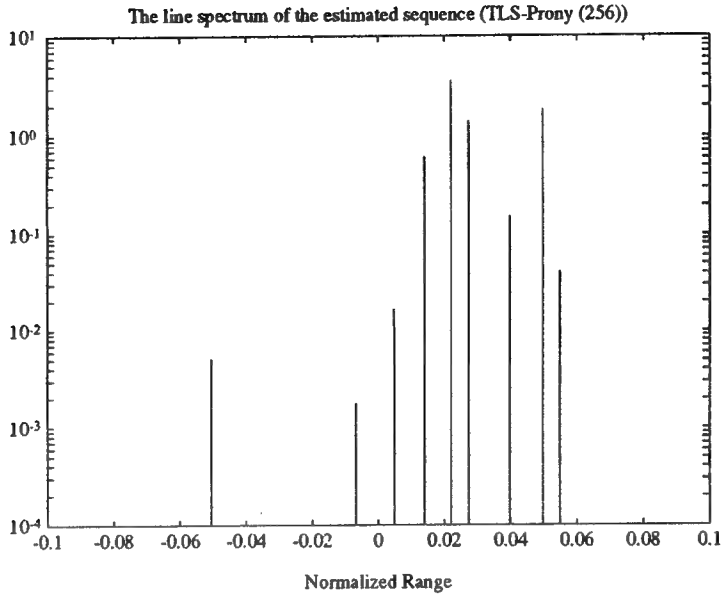


Figure 12: The line spectrum of the TLS-Prony estimates using the middle 256 points of the original data.

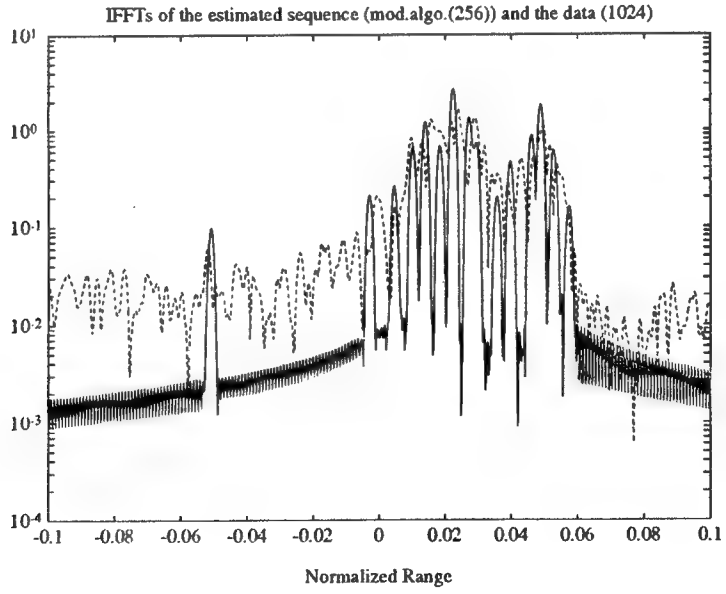


Figure 13: The IFFTs of the original data sequence and the estimated data sequence by the modified algorithm using the 256 points of the original data.

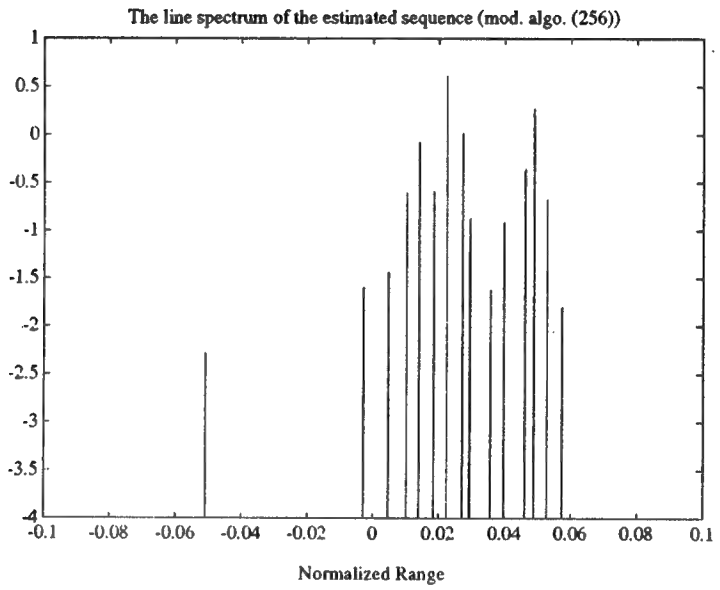


Figure 14: The line spectrum of the modified algorithm estimates using the middle 256 points of the original data.

IV. Conclusions

In this report we presented some preliminary results on parametric modeling of linear FM radar data collected at Rome Laboratories. Modeling techniques and the conventional FFT technique were compared. We focused on the TLS-Prony algorithm and proposed a modified algorithm incorporating TLS-Prony, a residue concept, and MLE. Several issues were discussed in the report. First, for the same bandwidth, the modeling techniques used in general outperform the FFT technique, especially the modified algorithm. In addition, we found that using the modified algorithm we may reduce the bandwidth by a factor of two without losing the performance in comparison to the FFT technique. Further investigation is needed on the modified algorithm in order to fully understand its performance.

References

[1] R. Carrière and R. L. Moses, "Autoregressive moving average modeling of radar target signatures," in *Proceedings of the IEEE 1988 National Radar Conference*, (Ann Arbor, MI), pp. 225–229, April 20–21, 1988.

[2] M. A. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP–35, no. 10, pp. 1440–1454, October 1987.

[3] W. M. Steedly, C. J. Ying, and R. L. Moses, "Statistical analysis of TLS-based Prony techniques," *Automatica*, Submitted for publication.

[4] P. Stoica, R. Moses, B. Friedlander, and T. Söderdtröm, "Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, no. 3, pp. 378–392, March 1989.

Rome Laboratory

Customer Satisfaction Survey

RL-TR-_____

Please complete this survey, and mail to RL/IMPS, 26 Electronic Pky, Griffiss AFB NY 13441-4514. Your assessment and feedback regarding this technical report will allow Rome Laboratory to have a vehicle to continuously improve our methods of research, publication, and customer satisfaction. Your assistance is greatly appreciated.
Thank You

_____

_____

Organization Name:_____(Optional)

Organization POC: _____(Optional)

Address:_____

1. On a scale of 1 to 5 how would you rate the technology developed under this research?

5-Extremely Useful 1-Not Useful/Wasteful

Rating_____

Please use the space below to comment on your rating. Please suggest improvements. Use the back of this sheet if necessary.

2. Do any specific areas of the report stand out as exceptional?

Yes___ No_____

If yes, please identify the area(s), and comment on what aspects make them "stand out."

3. Do any specific areas of the report stand out as inferior?

Yes___ No___

If yes, please identify the area(s), and comment on what aspects make them "stand out."

4. Please utilize the space below to comment on any other aspects of the report. Comments on both technical content and reporting format are desired.

*U.S. GOVERNMENT PRINTING OFFICE: 1995-610-126-50224

MISSION

OF

ROME LABORATORY

Mission. The mission of Rome Laboratory is to advance the science and technologies of command, control, communications and intelligence and to transition them into systems to meet customer needs. To achieve this, Rome Lab:

a. Conducts vigorous research, development and test programs in all applicable technologies;

b. Transitions technology to current and future systems to improve operational capability, readiness, and supportability;

c. Provides a full range of technical support to Air Force Materiel Command product centers and other Air Force organizations;

d. Promotes transfer of technology to the private sector;

e. Maintains leading edge technological expertise in the areas of surveillance, communications, command and control, intelligence, reliability science, electro-magnetic technology, photonics, signal processing, and computational science.

The thrust areas of technical competence include: Surveillance, Communications, Command and Control, Intelligence, Signal Processing, Computer Science and Technology, Electromagnetic Technology, Photonics and Reliability Sciences.